AD-770 067

MODEL ERROR COMPENSATION TECHNIQUES FOR LINEAR FILTERING

Hamilton Hagar, Jr.

Texas University

Prepared for:

Air Force Office of Scientific Research

August 1973

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD770067

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR - TR - 73 - 2040

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle): MODEL ERROR COMPENSATION TECHNIQUES FOR LINEAR FILTERING

5. TYPE OF REPORT & PERIOD COVERED: Interim

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): Hamilton Hagar

8. CONTRACT OR GRANT NUMBER(s): AFOSR 72-2233

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
University of Texas at Austin
Dept of Aerospace Engr & Engr Mechanics
Austin, Texas 78712

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
61102F
9769-01

11. CONTROLLING OFFICE NAME AND ADDRESS:
Air Force Office of Scientific Research/NM
1400 Wilson Blvd
Arlington, Virginia 22209

12. REPORT DATE: August 1973

13. NUMBER OF PAGES: 234 246

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

A. Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Orbit Determinates
Guidance and Control
Adaptive Estimation Methods
Low-Thrust
Unmodel Force Compensation

Reproduced by
NATIONAL TECHNICAL INFORMATION SERVICE
U S Department of Commerce
Springfield VA 22151

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The exceptional utility and performance of the sequential, linear, unbiased, minimum variance estimator suffers severely in the presence of dynamic model errors. This problem--perhaps the greatest detriment to the so-called Kalman filter algorithm-- is discussed in light of its divergent effect upon the estimation process.

(continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

1a

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

BLOCK #20/Abstract

A number of optimal and suboptimal modifying techniques are described which attempt to prevent this divergence. Extensions are developed resulting in adaptive forms and a new algorithm is derived for sequentially estimating the state noise covariance matrix. Performance of the techniques is illustrated by their application to, (1) the terminal phase of an Earth orbit rendezvous mission, and (2) the heliocentric trajectory determination of a solar electric propulsion space vehicle. Numerical results indicate that the model error difficulties can be sufficiently countered, with particularly effective performance being supplemented by the sequential state noise covariance estimator.

MODEL ERROR COMPENSATION TECHNIQUES FOR LINEAR FILTERING

Hamilton Hagar, Jr., B.S., M.S.M.E.
The University of Texas at Austin
Austin, Texas

AMRL 1055
August 1973

Applied Mechanics Research Laboratory
The University of Texas at Austin
Austin, Texas

This report was prepared under

Contract No. JPL 953147

for the

Jet Propulsion Laboratories

and

Grant No. AFOSR 72-2233

for the

Air Force Office of Scientific Research

by the

Applied Mechanics Research Laboratory
The University of Texas at Austin
Austin, Texas

under the direction of

Byron D. Tapley
Professor and Chairman

Preceding page blank

MODEL ERROR COMPENSATION TECHNIQUES FOR LINEAR FILTERING

Hamilton Hagar, Jr., Ph.D.
The University of Texas at Austin, 1973

Supervising Professor: Byron D. Tapley

The exceptional utility and performance of the sequential, linear, unbiased, minimum variance estimator suffers severely in the presence of dynamic model errors. This problem -- perhaps the greatest detriment to the so-called Kalman filter algorithm -- is discussed in light of its divergent effect upon the estimation process.

A number of optimal and suboptimal modifying techniques are described which attempt to prevent this divergence. Extensions are developed resulting in adaptive forms and a new algorithm is derived for sequentially estimating the state noise covariance matrix. Performance of the techniques is illustrated by their application to, (1) the terminal phase of an Earth orbit rendezvous mission, and (2) the heliocentric trajectory determination of a solar electric propulsion space vehicle. Numerical results indicate that the model error difficulties can be sufficiently countered, with particularly effective performance being supplemented by the sequential state noise covariance estimator.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Background and Scope

The first solution to the problem of optimally estimating the values of a set of quantities from a large set of data is generally attributed to Karl F. Gauss' method of least squares (1).* Although A. M. Legendre offered an early published version (2) in 1806, Gauss provided the basic mathematical derivation. Interestingly, the method was developed and applied to classical problems of orbit determination. Nearly 170 years later, it is now used as a fundamental technique in space vehicle tracking and modern orbit determination. In fact, the wide applicability and use of least squares in all fields of engineering is testimony to the genius and insight of Gauss.

Although intermittent developments of some importance occurred, particularly the ideas of probabilistic approaches, it was not until after the first decade of the twentieth century that the foundations of estimation theory were extended at a level of significance parallel to Gauss'. R. A. Fisher (3) introduced many of the terms used to characterize the performance of estimators; his concepts and efforts provided fertile ground for further developments and new approaches to estimation theory. In 1942, Norbert Wiener, considered today as one of the world's leading mathematical analysts, produced the so-called Wiener-Hopf integral equation. The solution to this

---

* Parenthesized numbers indicate references as enumerated in the Reference section. When specific pages are referenced, they are separated from the reference number by a comma. Thus (5, 10-12) indicates reference 5, pages 10 through 12.

equation is a weighting function which, when combined with a linear measurement, results in an estimate of the desired quantity which minimizes the error in a mean-square sense (more exactly the Wiener-Hopf integral offers a method for producing linear, minimum variance unbiased estimates).

Although Wiener's work is truly significant, the integral equation is limited in its practical application. With the growth of statistical communication theory, Wiener's technique received wide attention. Several attempts were made to improve and generalize the theory; however, none of these increased its basic utility and applicability.

The development of the digital computer provided a practical alternative approach to extending the applicability of estimation theory. Rather than attempt analytic extensions and solutions to the Wiener-Hopf integral equation, R. E. Kalman and R. S. Bucy (4) derived a differential equation from the Wiener-Hopf integral. The computational efficiency of the digital computer made the numerical solution of the differential equation practical and resulted in a widely applicable algorithm for providing linear, unbiased, minimum-variance estimates. Today, particularly in navigation and guidance applications, "Kalman Filtering" as it has come to be called, ranks next to least squares in popularity.

In spite of the utility of the Kalman filtering algorithm, the technique suffers from a particularly severe problem known as divergence of the estimate. It usually arises from the fact that for the state vector to be estimated, the system dynamic model is incorrect. Operation of the Kalman filtering algorithm in the presence of modeling errors produces estimates which are essentially worthless: the estimated state is grossly in error.

Divergence of the estimate, perhaps the greatest detriment to the Kalman filtering algorithm, has received considerable attention. As a result,

various techniques have been devised to compensate for modeling error. Such techniques may be thought of as falling loosely into two major categories: adaptive and non-adaptive methods. Non-adaptive methods generally attempt to improve the estimation process by altering the filter structure in an *a priori* and thus suboptimal manner. They also include *a priori* approximations to the actual modeling errors. Adaptive methods attempt to improve knowledge of the dynamic model or to improve operation of the filter *during* the estimation process.

The basic objective of this study is to investigate the utility of a variety of model error compensation techniques, both adaptive and non-adaptive, and to compare the effectiveness of these methods. In the remainder of Chapter 1, the Kalman filter is introduced and its properties discussed. The model error problem is illustrated by a simple example. A brief literature survey of model error compensation techniques is also presented. In Chapter 2, selected non-adaptive error compensation methods are presented and compared analytically. In Chapter 3, various new and previously developed adaptive methods are discussed. Chapters 4 and 5 illustrate the application of the various techniques to selected problems, and Chapter 6 concludes the study.

1.2 The Kalman Estimator and Some of Its Properties

Many expositions of the Kalman filtering algorithm exist in the literature, offering a number of unique, yet unifying approaches to the theory (5, 195-209). Thus our purpose here is not to give a rigorous derivation of the Kalman estimator. However, as a point of departure and for the sake of consistency, a somewhat heuristic development will be presented. Following this, some of the notable characteristics of the algorithm will be discussed.

Before proceeding, an important preliminary must be treated. For a given dynamic system, the problem is generally one of observing or measuring, in a stochastic environment, some aspect of operation of the system, and then treating the measurement output in a judicious manner in order to extract specific knowledge of the system's performance. However, an important distinction must be made concerning the measurement process. On the one hand is the case where the measurement process is carried out continuously in time; this is often found in analog computing applications. On the other hand is the case where measurements are made at discrete points in time in correspondence with digital computing applications. While formulations of the linear, unbiased, minimum variance estimator are known for both cases, widespread application of the digital computer focuses attention upon the discrete formulation. In the subsequent developments we will be concerned with this approach only.

Within this scope the problem may be stated as follows: (1) Given a dynamical system modeled by the linear difference equation

$$x_k = \Phi_{k,k-1}\, x_{k-1} + \Gamma_{k,k-1}\, w_{k-1}, \qquad (1.2.1)$$

where

$x_k = x(t_k)$ is an n-vector of random state variables, and $x_o = x(t_o)$ is given;

$\Phi_{k,k-1} = \Phi(t_k, t_{k-1})$ is the $n \times n$ state transition matrix*, with $\Phi_{j,j} = I$;

---

* The properties of the state transition matrix are well known. Some of these are mentioned later.

$w_k = w(t_k)$ is an m-vector of random input disturbances with the statistics

$$E\{w_k\} = 0,\ E\{w_k\, w_j^T\} = Q_k\, \delta_{kj},\ E\{x_{k-1}\, w_{k-1}^T\} = 0,$$

(δ_{kj} is the Kronecker delta; $(\ )^T$ denotes transpose);

$\Gamma_{k,k-1} = \Gamma(t_k, t_{k-1})$ is the $n \times m$ disturbance transition matrix; and

(2) Given the linear observation-state relationship

$$y_k = H_k\, x_k + v_k \qquad (1.2.2)$$

where

$y_k = y(t_k)$ is a p-vector of observations;

$H_k = H(t_k)$ is a $p \times n$ mapping matrix;

$v_k = v(t_k)$ is a random p-vector of observation noise with statistics

$$E\{v_k\} = 0,\ E\{v_k\, v_j^T\} = R_k\, \delta_{kj},\ E\{x_{k-1}\, v_k^T\} = 0,$$

$$E\{v_k\, w_j^T\} = 0, \text{ for all } k \text{ and } j \text{ *};$$

(3) Find an unbiased estimate, $\hat{x}_k$, based on the set, Y_k, of k observation vectors through time t_k,

$$Y_k \equiv \{y_1, y_2, \ldots, y_k\},$$

i.e., find $\hat{x}_k \equiv E\{x(t_k|Y_k)\}$ with $E\{\hat{x}_k\} = x_k$;

(4) Such that $\hat{x}_k$ is formed as a linear combination of the k observation vectors:

$$\hat{x}_k = L_k\, \hat{x}_{k-1} + K_k\, y_k, \qquad (1.2.3)$$

---

* If Q_k and R_k are constant for all k, the corresponding noise processes are wide-sense stationary.

and such that the state error covariance

$$P_k = P(t_k|Y_k) \equiv E\{(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T\} \tag{1.2.4}$$

at time, t_k, based on Y_k, is minimized. Note in *(1.2.3)* that $\hat{x}_{k-1}$ is the estimate of x_{k-1} based on Y_{k-1}, and in this sense $\hat{x}_k$ is based on all of the elements of Y_k. In other words, given *(1.2.1)* and *(1.2.2)*, the problem is to find L_k and K_k such that *(1.2.3)* is an unbiased, minimum variance estimate of x_k.

In order to aid the discussion, some notational simplifications are made. In general, the notation

$$f_{k|k} = f(t_k|Y_k)$$

means the function, f, evaluated at t_k, based on the knowledge of Y_k. For brevity we will simply equivalence this with the notation, f_k. When no confusion exists, the subscript k will be eliminated altogether. Hence,

$$\Phi = \Phi_{k,k-1} = \Phi(t_k, t_{k-1})$$

$$\Gamma = \Gamma_{k,k-1} = \Gamma(t_k, t_{k-1})$$

$$x = x_k = x(t_k)$$

$$\hat{x} = \hat{x}_k = \hat{x}_{k|k} = \hat{x}(t_k|Y_k)$$

$$\overline{x} = \overline{x}_k = \overline{x}_{k|k} = \Phi\,\hat{x}_{k-1}$$

$$\tilde{x} = \tilde{x}_k = \tilde{x}_{k|k} = x - \hat{x}$$

$$\overline{\tilde{x}} = \overline{\tilde{x}}_k = \tilde{x}_{k|k-1} = x - \overline{x}$$

$$Q = Q_k = Q(t_k)$$

$$R = R_k = R(t_k)$$

$$P = P_k = P_{k|k} = P(t_k|Y_k)$$

$$\overline{P} = \overline{P}_k = P_{k|k-1} = P(t_k|Y_{k-1}) = \Phi P_{k-1}\Phi^T + \Gamma Q_{k-1}\Gamma^T \tag{1.2.5}$$

$$L = L_k = L(t_k)$$

$$K = K_k = K(t_k)$$

Proceeding, the requirement which must be satisfied if $\hat{x}$ is unbiased can be obtained by using *(1.2.1)*, *(1.2.2)*, and *(1.2.3)* to form the state estimate error, $\tilde{x}$, and then taking the expected value. Thus,

$$\tilde{x}_k = x_k - \hat{x}_k = \Phi x_{k-1} + \Gamma \omega_{k-1} - L_k \hat{x}_{k-1} - K_k y_k$$

$$= \Phi x_{k-1} + \Gamma \omega_{k-1} - L_k \hat{x}_{k-1} - K_k (H_k \Phi x_{k-1} + H_k \Gamma \omega_{k-1} + v_k)$$

$$E\{\tilde{x}_k\} = x_k - E\{\hat{x}_k\} = 0$$

$$= \Phi x_{k-1} - L_k E\{\hat{x}_{k-1}\} - K_k H_k \Phi x_{k-1}$$

$$= [(I - KH)\Phi - L] x_{k-1} = 0 \qquad (1.2.6)$$

where use is made of the facts that $E\{\omega_{k-1}\} = 0$ and $E\{v_k\} = 0$. Satisfaction of equation *(1.2.6)* leads to the requirement that

$$L = (I - KH)\Phi \qquad (1.2.7)$$

and thus *(1.2.3)* becomes

$$\hat{x}_k = (I - KH)\Phi \hat{x}_{k-1} + K y_k$$

$$= \Phi \hat{x}_{k-1} + K(y_k - H \Phi \hat{x}_{k-1}) \qquad (1.2.8)$$

By using *(1.2.1)*, *(1.2.2)*, and *(1.2.8)* the state error and the state estimate error covariance matrix can be obtained as follows:

$$\tilde{x}_k = \Phi \tilde{x}_{k-1} + \Gamma \omega_{k-1} - K(H \Phi \tilde{x}_{k-1} + H \Gamma \omega_{k-1} + v)$$

$$= (I - KH)(\Phi \tilde{x}_{k-1} + \Gamma \omega_{k-1}) - K v_k \qquad (1.2.9)$$

$$P_k = E\{\tilde{x}_k \tilde{x}_k^T\} = E\{(I - KH)(\Phi \tilde{x}_{k-1} + \Gamma \omega_{k-1}) \cdot$$

$$(\tilde{x}_{k-1}^T \Phi^T + \omega_{k-1}^T \Gamma^T)(I - H^T K^T) - (I - KH)(\Phi \tilde{x}_{k-1} + \Gamma \omega_{k-1}) v_k^T K^T -$$

$$K v_k (\tilde{x}_{k-1}^T \Phi^T + \omega_{k-1}^T \Gamma^T)(I - H^T K^T) + K v_k v_k^T K^T\}$$

Combining terms, and recalling that $E\{x_{k-1}\, v_k^T\} = 0$, $E\{x_{k-1}\, w_k^T\} = 0$, $E\{w_k\, v_k^T\} = 0$ and $E\{v_k\, v_{k-1}^T\} = 0$, (see page 4) one obtains

$$P_k = (I - KH)\,(\Phi\, P_{k-1}\, \Phi^T + \Gamma\, Q_{k-1}\, \Gamma^T)(I - H^T K^T) + K\, R\, K^T \qquad (1.2.10)$$

In view of the definition of $\overline{P}$ in equation *(1.2.5)*,

$$P = (I - KH)\,\overline{P}\,(I - H^T K^T) + K\, R\, K^T \qquad (1.2.11)$$

The first variation is now taken with respect to K.

$$\delta P = (I - KH)\,\overline{P}\,(-H^T\, \delta K^T) + K\, R\, \delta K^T + (I - \delta K H)\,\overline{P}\,(I - H^T K^T) + \delta K\, R\, K^T \qquad (1.2.12)$$

Necessary and sufficient conditions that P be a minimum are that (1) $\delta P = 0$ and (2) that the second variation of P be positive definite. Solving *(1.2.12)* for K with $\delta P = 0$ leads to

$$K = \overline{P}\, H^T (H\, \overline{P}\, H^T + R)^{-1} \qquad (1.2.13)$$

For an arbitrary n-vector, s, the corresponding quadratic form for the second variation of P is required to be

$$s^T\, \delta^2 P\, s = s^T\, \delta K (H\, \overline{P}\, H^T + R)\, \delta K^T\, s > 0 \qquad (1.2.14)$$

Thus if P is positive definite, then the bracketed term in *(1.2.14)* must be positive definite, and fulfillment of *(1.2.14)* guarantees the existence of the inverse in *(1.2.13)*.

With K given by *(1.2.13)* the updated state estimate error covariance may be obtained by substituting for K in equation *(1.2.11)*.

$$P = (I - KH)\,\overline{P}\,(I - H^T K^T) + K\, R\, K^T$$

$$= \overline{P} - K\, H\, \overline{P} - \overline{P}\, H^T K^T + K (H\, \overline{P}\, H^T + R)\, K^T$$

$$= \overline{P} - K\, H\, \overline{P} - \overline{P}\, H^T K^T + \overline{P}\, H^T (H^T\, \overline{P}\, H^T + R)^{-1}\, (H\, \overline{P}\, H^T + R)\, K^T$$

or

$$P = (I - KH)\,\overline{P} \qquad (1.2.15)$$

which is the required covariance update equation.

The system dynamics are given exactly by equation *(1.2.1)*. The random input, w, is unknown; however, being probabilistic in nature, the expected value of the equation may be taken conditioned upon the observations up through time, t_k. Thus

$$E\{x_k|Y_{k-1}\} = \Phi_{k,k-1}\, E\{x_{k-1}|Y_{k-1}\} + \Gamma_{k,k-1}\, E\{w_{k-1}|Y_{k-1}\} ,$$

with

$$E\{w_k|Y_k\} = E\{w_k\} = 0$$

or

$$\overline{x}_k = \hat{x}(t_k|Y_{k-1}) = \Phi_{k,k-1}\, \hat{x}_{k-1} = \Phi_{k-1,k}\, \hat{x} \quad . \qquad (1.2.16)$$

Equation *(1.2.16)* provides the estimate of x at time t_k based upon observations through time t_{k-1}. It is also an unbiased, minimum variance estimate (5, 201). and

$$\overline{P}_k = P(t_k|Y_{k-1}) = \Phi\, P_{k-1}\, \Phi^T + \Gamma\, Q\, \Gamma^T \qquad (1.2.17)$$

is the predicted value of the error covariance at t_k given observations through time t_{k-1} (5, 201). Equation *(1.2.17)* follows directly by forming the state estimate error covariance using equations *(1.2.1)* and *(1.2.16)*.

Equations *(1.2.16)*, *(1.2.17)*, *(1.2.13)*, *(1.2.8)* and *(1.2.15)* are the equations of the classical linear, unbiased, minimum variance estimate. The algorithm, denoted as Algorithm I, is summarized below.

Given: the *a priori* information $\hat{x}_0$, P_0, y_k, and $k = 1$. Compute:

1. $\overline{x}_k = \Phi\, \hat{x}_{k-1}$
2. $\overline{P}_k = \Phi\, P_{k-1}\, \Phi^T + \Gamma\, Q_{k-1}\, \Gamma^T$
3. $K_k = \overline{P}_k\, H_k (H_k\, \overline{P}_k\, H_k^T + R_k)^{-1}$
4. $\hat{x}_k = \overline{x}_k + K_k(y_k - H_k\, \overline{x}_k)$
5. $P_k = (I - K_k\, H_k)\, \overline{P}_k$
6. $k = k + 1$, *repeat*

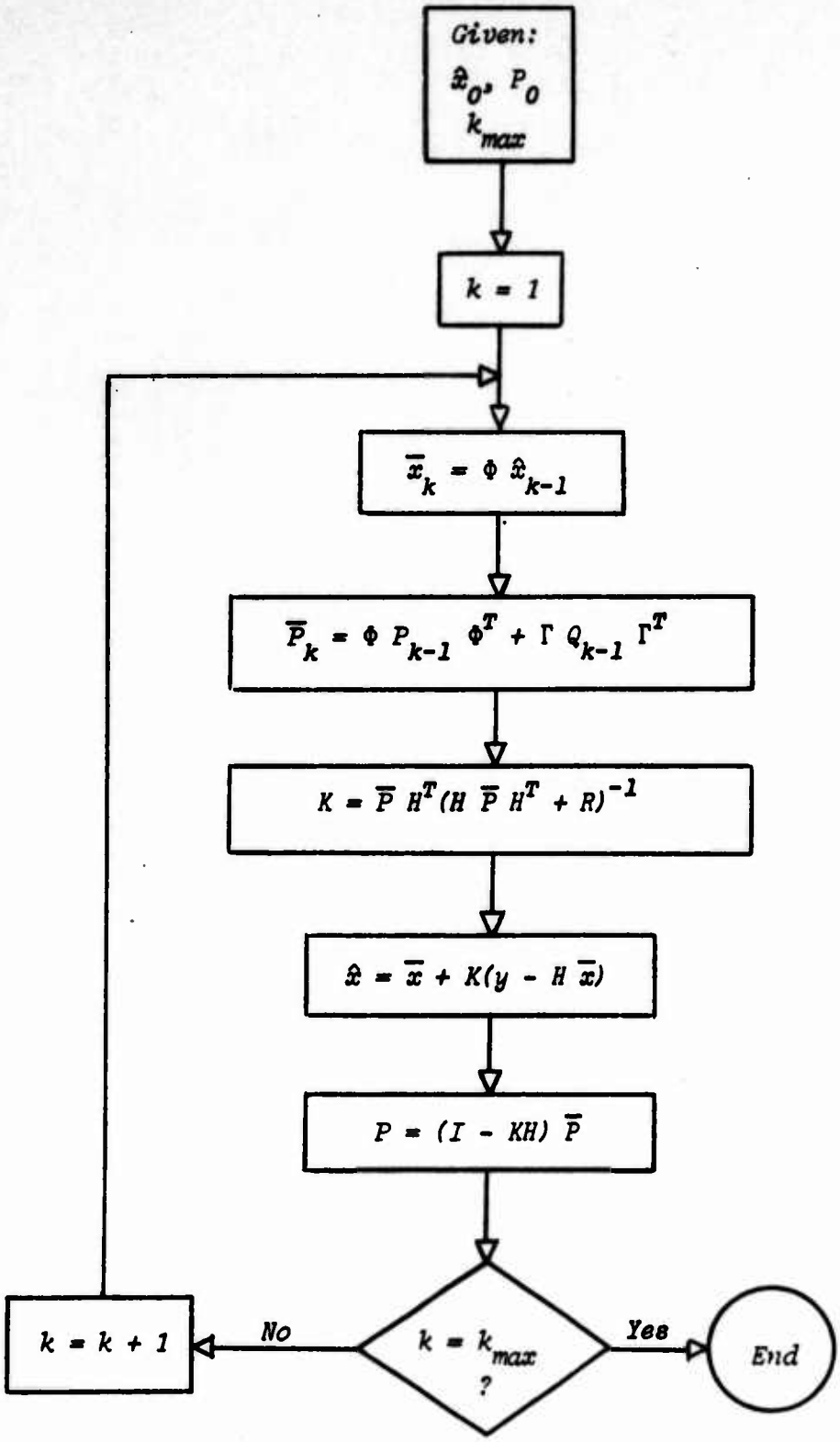


Figure 1.1 Kalman Filter, Algorithm I

The Kalman estimator above is based on linear system dynamics. It has been applied successfully to many nonlinear problems in the following manner.

Consider the n-element state vector, $X(t)$, whose values at discrete times are given by the nonlinear difference equation,

$$X(t_k) = F[X(t_{k-1}),\ t_k] + \Gamma_{k,k-1}\ w(t_{k-1}), \tag{1.2.18}$$

where w is a random input with properties as specified by (1.2.1), and Γ is the corresponding disturbance transition matrix. Define a reference state, $X^*(t)$. If $F[X(t_{k-1}),\ t_i]$ is continuous for $t_k \geq t_i \geq t_{k-1}$, then a linear approximation to F may be obtained by expanding in a Taylor series about $X^*(t_{k-1})$ to obtain

$$F(X_{k-1},\ t_k) \simeq F(X_{k-1}^*, t_k) + \frac{\partial F_k^*}{\partial X_{k-1}} (X_{k-1} - X_{k-1}^*)$$

$$\simeq X_k^* + \frac{\partial F_k^*}{\partial X_{k-1}} (X_{k-1} - X_{k-1}^*), \tag{1.2.19}$$

where the subscripts denote values at the corresponding times. Substituting into (1.2.18) we have, to first order in X_{k-1}.

$$X_k = X_k^* + \frac{\partial F_k^*}{\partial X_{k-1}} (X_{k-1} - X_{k-1}^*) + \Gamma_{k,k-1}\ w_{k-1}$$

Defining $x_j \equiv X_j - X_j^*$ and $\Phi_{k,k-1} \equiv \partial F_k^*/\partial X_{k-1}$†, then obtain (1.2.1).

$$x_k = \Phi_{k,k-1}\ x_{k-1} + \Gamma_{k,k-1}\ w_{k-1} \tag{1.2.1}$$

A similar linearization is performed if the observation-state relation is also nonlinear, but with additive noise, $v(t_k)$.

$$Y(t_k) = G[X(t_k),\ t_k] + v(t_k) \tag{1.2.20}$$

---

† Note that this is the definition of Φ for the linear case as well.

where $Y(t_k)$ is a p-vector of observations. Defining a nominal or reference observation,

$$Y_k^* = G[X_k^*, t_k] + v_k, \tag{1.2.21}$$

where the subscripts again denote values at the corresponding time, and expanding in a Taylor series we have

$$Y_k = Y_k^* + \frac{\partial G_k^*}{\partial X_k}(X_k - X_k^*) + \ldots + v_k, \tag{1.2.22}$$

or to first order in X_k,

$$y_k = H_k x_k + v_k \tag{1.2.23}$$

where

$$y_k = Y_k - Y_k^*,$$

$$H_k = \partial G_k^*/\partial X_k$$

Thus the optimal linear estimator previously presented may be applied to the cases of nonlinear dynamics and observation processes. However, accurate estimates of the deviation, $\hat{x}_k$, from the reference or nominal solution require that the linearity assumption be valid. It is possible that large deviations can violate the linearity assumption resulting in divergence of the estimate. One method of reducing this possibility is to employ the so-called extended Kalman filter in which a new reference is chosen at each observation. In particular, after each observation is processed, the reference solution is updated by setting the nominal equal to the new estimate:

$$X_k^* = \hat{X}_k = F(X_{k-1}^*, t_k) + \hat{x}_k \tag{1.2.24}$$

After forming X_k^*, $\hat{x}_k$ is set to zero so that $\hat{x}_k = \bar{x}_{k+1} = 0$. Then the next updated estimate, at t_{k+1}, is found as

$$\hat{x}_{k+1} = K_{k+1} y_{k+1}$$

or, in terms of X,

$$\hat{X}_{k+1} = F(X^*_k,\ t_{k+1}) + K_{k+1}\, y_{k+1} \qquad (1.2.25)$$

The notation X^*_k is no longer necessary since it is identical to $\hat{X}_k$. The extended Kalman algorithm, Algorithm II, is summarized in Figure 1.2 for the nonlinear problem.

Usually the difference equation *(1.2.1)* is obtained from the solution to the linear vector differential equation

$$\dot{x}(t) = A(t)\, x(t) + B(t)\, u(t) \qquad (1.2.27)$$

where $u(t)$ is some input. For the nonlinear problem, $A(t) = \partial f(X^*,\ t)/\partial X^*(t)$, where X^* satisfies the nonlinear differential equation, $\dot{X}^*(t) = f(X^*, t)$. The solution to *(1.2.27)* is well known (39, 41-43) (15, 31-43) in the form

$$x(t_k) = \Phi(t_k,\ t_{k-1})\, x(t_{k-1}) + \int_{t_{k-1}}^{t_k} \Phi(t_k,\ \tau)\, B(\tau)\, u(\tau)\, d\tau \qquad (1.2.28)$$

The state transition matrix, Φ, satisfies the homogeneous differential equation

$$\dot{\Phi}(t,\ t_j) = A(t)\, \Phi(t,\ t_j),\ \Phi(t_j,\ t_j) = I \qquad (1.2.29)$$

Φ also has the following properties:

$$\Phi(t_i,\ t_k) = \Phi(t_i,\ t_j)\, \Phi(t_j,\ t_k)$$
$$\Phi^{-1}(t_i,\ t_j) = \Phi(t_j,\ t_i)$$

When $u(t)$ is replaced by white noise, with $E\{u(t)\, u^T(\tau)\} = U(t)\, \delta(t-\tau)$[†], difficulties occur in attempting to evaluate the resulting stochastic integral in *(1.2.28)*. The problem arises from the fact that the elements of white noise are uncorrelated in time, and hence are nowhere continuous. A common approach to remedy this (44, 326-327) (39, 115-117) is to select $\Delta t \equiv t_k - t_{k-1}$ small compared to the system characteristic

---

[†] $\delta(t - \tau)$ is the Dirac delta.

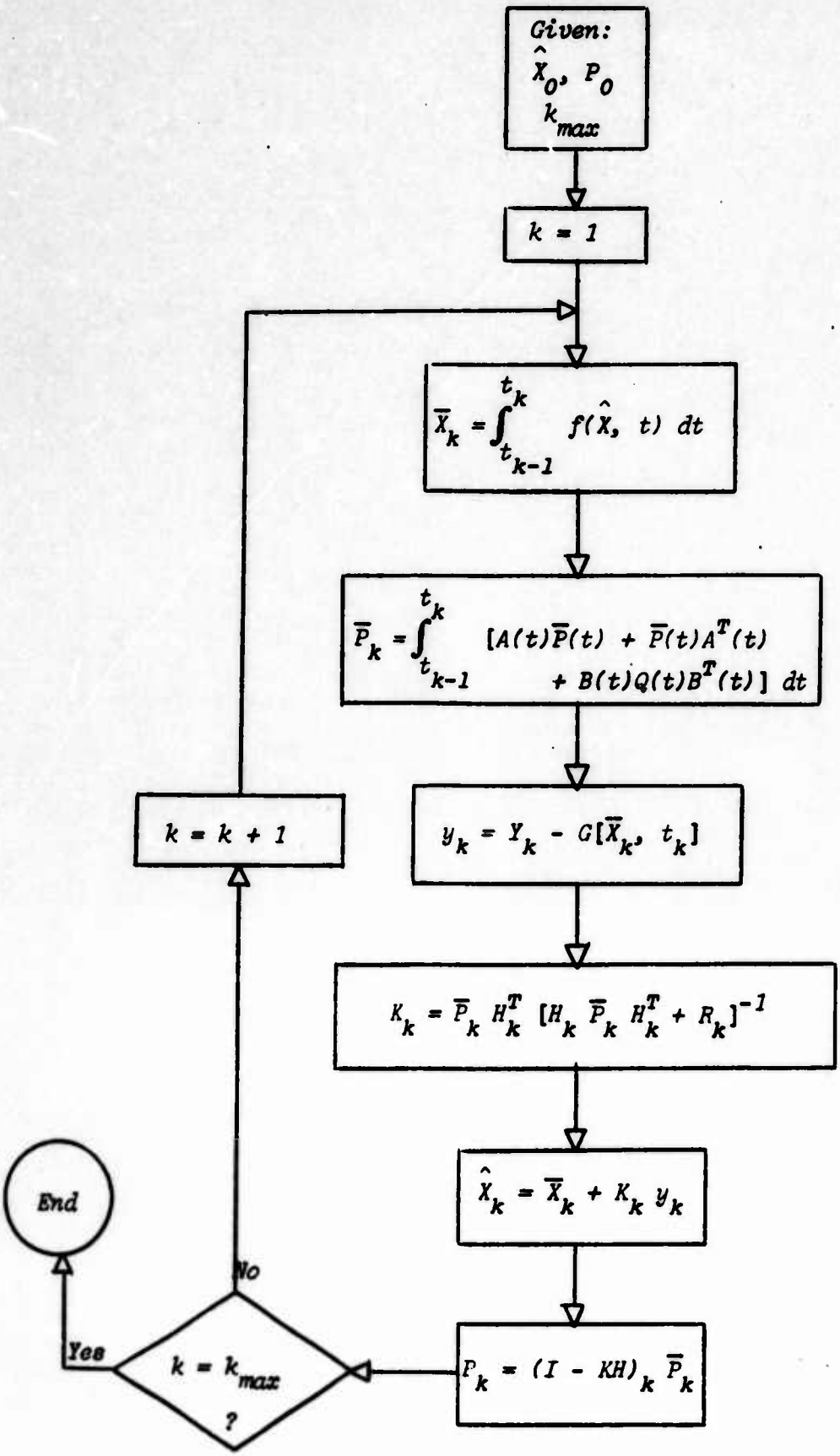


Figure 1.2 Extended Kalman Filter: Algorithm II

response time, and approximate $u(\tau)$ as a constant $w(t_{k-1})$ over Δt.*

In this case we define

$$\Gamma(t_k, t_{k-1}) \equiv \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau)\, B(\tau)\, d\tau \qquad (1.2.30)$$

This is the approach taken to obtain equation *(1.2.1)*. An alternative method is to define, formally,

$$d\omega' = u(\tau)\, d\tau$$

and employ the definition of the Itô stochastic integral (5, 98) to obtain

$$\int_{t_{k-1}}^{t_k} \Phi(t_k, \tau)\, B(\tau)\, d\omega = \Phi_{k,k-1}\, B_k\, (\omega'_k - \omega'_{k-1}) \qquad (1.2.31)$$

In this case one cannot write the form $\Gamma_{k,k-1}\, w_{k-1}$. Another alternative, the Stratonovich stochastic integral, yields the same result for this case (5, 116-120). The most general consideration is simply to define

$$w(t_{k-1}) \equiv \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau)\, B(\tau)\, u(\tau)\, d\tau \qquad (1.2.32)$$

In this case, $\Gamma_{k,k-1} = I$. The differences in these definitions are manifested in the estimation process through the corresponding covariance terms. Thus for *(1.2.30)* one finds

$$\Gamma\, E\{w\, w^T\}\, \Gamma^T =$$

$$\int_{t_{k-1}}^{t_k} \Phi(t_k, \tau)\, B(\tau)\, d\tau\; E\{u_{k-1}\, u_{k-1}^T\} \int_{t_{k-1}}^{t_k} B^T(\tau)\, \Phi^T(t_k, \tau)\, d\tau \qquad (1.2.33)$$

For the Itô integral,

$$\Gamma\, E\{w\, w^T\}\, \Gamma^T = \Phi_{k,k-1}\, B_k\, E\{(\omega'_k - \omega'_{k-1})(\omega'_k - \omega'_{k-1})\}\, B_k^T\, \Phi_{k,k-1}^T$$

Carrying out the multiplication, this becomes

---

* A white Gaussian process may be shown to be the limit of a white Gaussian sequence, e.g., (5, 83-85).

$$\Gamma E\{w\ w^T\}\ \Gamma^T = \Phi\ B\ E\{w_k'\ w_k'^T + w_{k-1}'\ w_{k-1}'^T\}\ B^T\ \Phi^T \qquad (1.2.34)$$

since $E\{w_i'\ w_j'^T\} = Q_i'\ \delta_{ij}$

Finally, for (1.2.32)

$$\Gamma E\{w\ w^T\}\ \Gamma^T = E\{w\ w^T\} =$$

$$\int_{t_{k-1}}^{t_k}\int \Phi(t_k,\ \tau)\ B(\tau)\ E\{u(\tau)\ u^T(\rho)\}\ B^T(\rho)\ \Phi^T(t_k,\ \rho)\ d\tau d\rho$$

But $E\{u(\tau)\ u^T(\rho)\} = U(\tau)\ \delta(\tau - \rho)$, and since (44, 332)

$$\int_{t_{k-1}}^{t_k} \delta(\tau - \rho)\ d\rho = \begin{cases} 1, & t_{k-1} < \tau < t_k \\ 0, & t_{k-1} > \tau > t_k \end{cases}$$

then

$$E\{w\ w^T\} = \int_{t_{k-1}}^{t_k} \Phi(t_k,\ \tau)\ B(\tau)\ U(\tau)\ B^T(\tau)\ \Phi(t_k,\ \tau)\ d\tau \qquad (1.2.35)$$

Noting $Q = E\{w_j\ w_j^T\} = E\{w\ w^T\}$, one finds for (1.2.33), (1.2.34), and (1.2.35), respectively,

$$\Gamma\ Q\ \Gamma^T = \Gamma\ U\ \Gamma^T \qquad (1.2.33a)$$

$$\Gamma\ Q\ \Gamma^T = 2\ \Phi\ B\ Q'\ B^T\ \Phi^T \qquad (1.2.34a)$$

$$\Gamma\ Q\ \Gamma^T = Q = \int_{t_{k-1}}^{t_k} \Phi\ B\ U\ B^T\ \Phi^T\ d\tau \qquad (1.2.35a)$$

Thus, depending upon how one *defines* the stochastic integral in (1.2.28) various forms may be obtained for $\Gamma\ Q\ \Gamma^T$.

With the extended Kalman form, the state transition matrix is used only in the prediction of the state error covariance matrix. It is possible to eliminate the state transition matrix altogether by use of a differential equation for the error covariance, obtained by a straightforward limiting

process. First we write *(1.2.17)* such that the forcing term $\Gamma Q \Gamma^T$ is in the form *(1.2.35)*. Letting $\tau = t_k$, $t = t_{k-1}$, and defining $\Delta t \equiv \tau - t$, we have

$$\bar{P}(\tau) = \Phi(\tau, t)\, P(t)\, \Phi^T(\tau, t) + \int_t^\tau \Phi(\tau, s)\, B(s)\, U(s)\, B^T(s)\, \Phi^T(\tau, s)\, ds \tag{1.2.36}$$

Using the Taylor series expansion,

$$\bar{P}(\tau) = [\Phi(t,t) + \dot{\Phi}(t,t)\, \Delta t + \ldots]\, P(t)\, [\Phi(t,t) + \dot{\Phi}(t,t)\, \Delta t + \ldots]^T$$

$$\left[\int_t^{\tau=t} \Phi(\tau,s)\, B(s)\, U(s)\, B^T(s)\, \Phi^T(\tau,s)\, ds + \right.$$

$$\left. \frac{d}{d\tau} \int_t^{\tau=t} \Phi(\tau,s)\, B(s)\, U(s)\, B^T(s)\, \Phi^T(\tau,s)\, ds\, \Delta t + \ldots\right] \tag{1.2.37}$$

where $\dot{\Phi}(t,t) = d\Phi(\tau,t)/d\tau$ evaluated at $\tau = t$. The first integral in *(1.2.37)* is zero. Applying Leibnitz' rule to the second integral we have

$$\frac{d}{d\tau} \int_t^{\tau=t} \Phi(\tau,s)\, B(s)\, U(s)\, B^T(s)\, \Phi^T(\tau,s)\, ds\, \Delta t =$$

$$B(t)\, U(t)\, B^T(t)\, \Delta t + \int_t^{\tau=t} \frac{\partial}{\partial \tau} [\Phi(\tau,s)\, B(s)\, U(s)\, B^T(s)\, \Phi^T(\tau,s)]\, ds\, \Delta t \tag{1.2.38}$$

Noting that the integral here is also zero, and that $\dot{\Phi}(t,t) = A(t)$, one finds on substituting into *(1.2.37)* that

$$\bar{P}(\tau) = P(t) + A(t)\, P(t)\, \Delta t + P(t)\, A^T(t)\, \Delta t + B(t)\, U(t)\, B^T(t)\, \Delta t + \ldots$$

Now, by definition $P(t_{k-1}) = P(t_{k-1}|Y_{k-1})$, and $\bar{P}(t_k) = P(t_k|Y_{k-1})$. Since no new information is being added, then at $t = t_{k-1}$ we may make the substitution $\bar{P}(t)$ for $P(t)$. Then substracting $\bar{P}(t)$ from both sides, dividing by Δt and taking the limit as $\Delta t \to 0$, one obtains the differential equation

$$\dot{\bar{P}}(t) = A(t)\, \bar{P}(t) + \bar{P}(t)\, A^T(t) + B(t)\, Q(t)\, B^T(t) \tag{1.2.39}$$

Since $\bar{P}$ is a symmetric $n \times n$ matrix, only $n(n+1)/2$ equations need to be integrated. Φ is $n \times n$ and, in general, not symmetric, so a reduction in numerical effort is obtained by integrating the appropriate $n(n+1)/2$ equations for $\dot{\bar{P}}$. On the other hand, depending upon the specific problem at hand, it is often possible to reduce the number of equations in $\dot{\Phi} = A\Phi$ by obtaining closed form solutions for some of the elements. Also, it has been found that *(1.2.39)* can be a difficult equation to integrate numerically.

1.3 Properties of the State Error Covariance

This section deals with some characteristics of the state error covariance matrix. We will not elaborate at length -- a number of investigators have done extensive work in this area, notably Kalman. However, some of the key properties are presented.

We first show that any covariance matrix is non-negative definite. For a vector random variable, z, with mean, μ, and covariance, Z, consider the quadratic form

$$q = a^T Z a \qquad (1.3.1)$$

where a is an arbitrary vector of constants. Since the expected value is a linear operator,

$$q = a^T E\{(z-\mu)(z-\mu)^T\}\, a = E\{a^T(z-\mu)(z-\mu)^T a\} \qquad (1.3.2)$$

Defining the scalar

$$s \equiv a^T(z-\mu)$$

we have

$$q = E\{s\, s^T\} = E\{s^2\} \qquad (1.3.3)$$

which is never negative. Thus Z is non-negative definite.

Next, positive definite bounds may be established for the error covariance associated with the Kalman filter. In particular, Sorenson (7)

takes a rather unique approach by decomposing the system, equations *(1.2.1)* and *(1.2.2)* into two separate parts. One, denoted as p, has input noise, but perfect measurements; the other, denoted as m, has no input noise; and imperfect measurements. Thus,

$$x_k = x_k^m + x_k^p$$

$$y_k = y_k^m + y_k^p \qquad (1.3.4)$$

where the systems m and p are

$$m: \begin{cases} x_p^m = \Phi x_{k-1}^m \\ y_p^m = H x_k^m + v_k \end{cases} \qquad (1.3.5)$$

$$p: \begin{cases} x_k^p = \Phi x_{k-1}^p + \Gamma w_{k-1} \\ y_k^p = H x_k^p \end{cases} \qquad (1.3.6)$$

Sorensen then discusses the covariance properties for these two systems separately. For the noise-free plant, m, the observability matrix, defined as

$$M_{k,j} = \sum_{i=j}^{k} \Phi_{i,j-1}^T H_i^T R_i^{-1} H_i \Phi_{i,j-1} \qquad (1.3.7)$$

is used to establish the positive definite property of the state error covariance matrix. Specifically, a matrix inversion lemma, the so-called Schur identity given in Appendix A allows *(1.2.15)* to be written as

$$P_k^m = [I - K_k H_k] \bar{P}_k^m = [\bar{P}_k^{-1m} + H_k^T R_k^{-1} H_k]^{-1} \qquad (1.3.8)$$

where now $\bar{P}_k^m = \Phi P_{k-1}^m \Phi^T$, since m is the noise free plant. Then,

$$P_k^{-1m} = \Phi_{k,k-1}^{-1} P_{k-1}^{-1m} \Phi_{k,k-1}^{-T} + H_k^T R_k^{-1} H_k \qquad (1.3.9)$$

The system is said to be observable on the interval $[t_j, t_k]$ if the observatility matrix $M_{k,j}$, is positive definite. If this condition holds, and P_0^m is also positive definite, then *(1.3.9)* can be written as

$$P_k^m = \Phi_{k,0} \, [P_0^{-1m} + M_{k,1}]^{-1} \, \Phi_{k,0}^T \qquad (1.3.10)$$

where use is made of the fact that

$$\begin{aligned} \overline{P}_k^{-1m} &= [\Phi_{k,k-1} \, P_{k-1}^m \, \Phi_{k,k-1}^T]^{-1} \\ &= \Phi_{k,k-1}^{-1} \, P_{k-1}^{-1m} \, \Phi_{k,k-1}^{-T} \\ &= \Phi_{k,k-1} \, P_{k-1}^{-1m} \, \Phi_{k,k-1}^T \end{aligned}$$

By arguments similar to the previous discussion, P_k^m is therefore positive definite for all k, since the bracketed term in *(1.3.10)* is also. Additionally, Sorenson shows that M may be described in a recursive fashion by

$$M_{k,j} = M_{k-1,j} + \Phi_{k,j-1}^T \, H_k^T \, R_k^{-1} \, H_k \, \Phi_{k,j-1} \qquad (1.3.11)$$

and since the second term on the right is non-negative definite, $M_{k,j} - M_{k-1,j}$ is also, thus allowing the conclusion, by *(1.3.10)*, that P_k^m is positive definite for all $k > j$.

Building on these facts, Sorenson also states the well known characteristic that P_k^m generally tends to vanish for well-defined systems with no input noise. First, the concept of q-stage observability is defined. Given $1 \leq q \leq N$ such that $t_1 \leq t_{k-q+1}$ and $t_k \leq t_N$, the system m is said to be q-stage observable on an interval $[t_0, t_N]$ if and only if $M_{k,k-q+1}$ is positive definite for every t_k. That this implies $M_{k,1} - M_{k-q,1}$ is positive definite may be readily seen by modifying *(1.3.11)* to obtain

$$M_{k,1} = M_{k-q,1} + \Phi_{k-q,0}^T \, M_{k,k-q+1} \, \Phi_{k-q,0} \qquad (1.3.12)$$

This follows by substituting $M_{k,k-q+1}$ using *(1.3.7)*. Under q-stage observability, *(1.3.12)* suggests that in some sense $M_{k,j}$ increases without bound. Consider the spectral norm of M and M^{-1}:

$$||M_{k,1}|| = \lambda^k_{max}, \quad ||M^{-1}_{k,1}|| = 1/\lambda^k_{min} \tag{1.3.13}$$

where λ^k_{max} and λ^k_{min} are the maximum and minimum eigenvalues of $M_{k,1}$. Now if $M_{k,1}$ is q-stage observable,

$$\lambda^k_{max} > \lambda^{k-q}_{max}, \quad \lambda^k_{min} > \lambda^{k-q}_{min},$$

with similar inequalities for the other eigenvalues. Then

$$\lim_{k\to\infty} ||M_{k,1}|| = \lim_{k\to\infty} \lambda^k_{max} = \infty,$$

and

$$\lim_{k\to\infty} ||M^{-1}_{k,1}|| = \lim_{k\to\infty} \frac{1}{\lambda^k_{min}} = 0 \tag{1.3.14}$$

Using this result together with *(1.3.10)*, Sorenson realizes the following conclusion: For a q-stage observable system, if $||M^{-1}_{k,1}||$ converges to zero more rapidly than $||\Phi_{k,0}||^2$ increases, then the error covariance P^m_k vanishes as $k \to \infty$. To verify this we merely need to show that the norm of the elements of *(1.3.10)* vanish. First, for no *a priori* information, $P^{-1m}_0 = 0$. Thus*,

$$\Phi_{k,0}\, M^{-1}_{k,1}\, \Phi^T_{k,0} > \Phi_{k,0}\, [P^{-1m}_0 + M_{k,1}]^{-1}\, \Phi^T_{k,0} \tag{1.3.15}$$

and

$$||P^m_k|| < ||\Phi_{k,0}\, M^{-1}_{k,1}\, \Phi^T_{k,0}|| \leq ||\Phi_{k,0}||^2\, ||M^{-1}_{k,1}|| \tag{1.3.16}$$

which vanish as $k \to \infty$ if the original hypotheses are met. Further, as $k \to \infty$, *(1.3.15)* tends to equality with increasing accuracy. The implication is that

---

* $A \geq B > C$ means that $A - B$ is non-negative definite and $B - C$ is positive definite.

for large k, P_k^m is essentially independent of P_0^m for q-stage observable systems.

For system p, where the measurements are perfect, one is able to conclude that the error covariance matrix is never positive definite. Using *(1.2.15)* and *(1.2.13)*,

$$H_k P_k^p = H_k[I - K_k H_k] \bar{P}_k^p$$

$$= \{H - H \bar{P}_k^p H^T[H \bar{P}_k^p H^T]^{-1} H\} \bar{P}_k^p$$

$$= 0 \qquad (1.3.17)$$

Thus if P_k^p were positive definite, H_k would have to assume the contradictory state of having to be identically zero. Utilizing the gains, K_k^m and K_k^p which are optimal for each system, Sorenson then shows that a lower bound on the total error covariance is given by

$$P_k \geq P_k^m + P_k^p$$

or

$$P_k \geq \Phi_{k,0}[P_0^{-1m} + M_{k,1}]^{-1} \Phi_{k,0} + P_k^p \qquad (1.3.18)$$

Utilizing the gain K_k^s, which is suboptimal for the total system but may be optimal for either (but not both) of the systems m and p, an upper bound is established as

$$P_k \leq P_k^{ms} + P_k^{ps}$$

where P^{ms} and P^{ps} are the corresponding covariances. If K_k^s is chosen to be K_k^m (the optimal gain for system m), and used for both systems, then

$$P_k \leq \Phi_{k,0}[P_0^{-1m} + M_{k,1}]^{-1} \Phi_{k,0} + P_k^{ps} \qquad (1.3.19)$$

But P_k^{pm} will generally increase without bound since noise is continually being input to the system, and therefore, as Sorenson points out, *(1.3.19)*

is an unsatisfactory bound. Note that in this case *(1.3.17)* generally will not hold, since the gain is not the optimal for system p. If in fact, $\Gamma Q_k \Gamma^T$ is positive definite, then

$$\begin{aligned} P_k^{pm} &= [I - \bar{P}^m H^T (H \bar{P}^m H^T + R)^{-1} H] \bar{P}_k^{pm} \\ &= (I - K^m H)[\Phi P_{k-1}^{pm} \Phi^T + \Gamma Q_{k-1} \Gamma^T] \\ &= (I - K^m H)[\bar{P}_k^m + \Delta \bar{P}^p + \Gamma Q_{k-1} \Gamma^T] \\ &\geq \Gamma Q_{k-1} \Gamma^T + \Delta \bar{P}^p \geq \Gamma Q_{k-1} \Gamma^T \end{aligned} \qquad (1.3.20)$$

This last is true since P_k^m and hence K^m eventually vanish. Thus *(1.3.18)* and *(1.3.19)* represent positive definite bounds upon the total system error covariance, P_k.

An alternate definition of observability requires that the information matrix, $T_{k,j}$, be positive definite:

$$T_{k,j} \equiv \sum_{i=j}^{k} \Phi_{i,k}^T H_i^T R_i^{-1} H_i \Phi_{i,k} > 0, \quad k \geq j \qquad (1.3.21)$$

If this condition holds, the system *(1.3.1, 1.3.2)* is said to be completely observable with respect to $\{y_j, y_{j+1}, \ldots, y_k\}$. Note that $T_{k,j}$ is related to the observability matrix by

$$M_{k,j} = \Phi_{k,j-1}^T T_{k,j} \Phi_{k,j-1}, \qquad (1.3.22)$$

and related to the error covariance by

$$P_k^{-1m} = \Phi_{0,k}^T P_0^{-1m} \Phi_{0,k} + T_{k,1} \qquad (1.3.23)$$

This latter follows from *(1.3.10)*. A recursion for T may also be obtained using *(1.3.11)* and *(1.3.22)*:

$$T_{k,j} = \Phi^T_{j-1,k} M_{k,j} \Phi_{j-1} = \Phi^T_{j-1,k} M_{k-1,j} \Phi_{j-1,k} + H^T_k R^{-1}_k H_k$$

$$= \Phi^T_{k-1,k} \Phi^T_{j-1,k-1} M_{k-1,j} \Phi_{j-1,k-1} \Phi_{k-1,k} + H^T_k R^{-1}_k H_k$$

or

$$T_{k,j} = \Phi^T_{k-1,k} T_{k-1,j} \Phi_{k-1,k} + H^T_k R^{-1}_k H_k, \quad k > j \tag{1.3.24}$$

A concept dual to that of observability is controllability. The controllability matrix is defined as

$$\Lambda_{k,0} \equiv \sum_{i=1}^{k} \Phi_{k,i} \Gamma_{i,i-1} Q_{i-1} \Gamma^T_{i,i-1} \Phi^T_{k,i} \tag{1.3.25}$$

The corresponding dynamic system is said to be completely controllable if, and only if, $\Lambda_{k,0} > 0$ for $k > 0$.

Jazwinski (5, 234-243), drawing upon the work of Kalman, Sorenson, and others, gives a comprehensive discussion of bounds and stability of the filter equations. Extending the concepts of observability and controllability, Jazwinski defines the system *(1.2.1, 1.2.2)* as being uniformly completely observable if there exist a positive integer, N, and positive constants, α and β, such that

$$0 < \alpha I \leq T_{k,k-N} \leq \beta I$$

for all $k > N$. Similarly the system is uniformly completely controllable if

$$0 < \gamma I \leq \Lambda_{k,k-N} \leq \delta I$$

where γ and δ are positive constants. Using these definitions, Jazwinski's results are summarized in the following:

Lemma. If the dynamical system *(1.2.1, 1.2.2)* is uniformly completely observable and uniformly completely controllable, and if $P_0 \geq 0$, then

a) P_k is uniformly bounded from above,

b) P_k is uniformly bounded from below,

c) $P_k > 0$,

for all $k \geq N$. Further, the filter of Algorithm I is asymptotically stable; that is,

$$||\Phi(t_k, t_0)|| \leq c_1 \exp[-c_2(t_k - t_0)]$$

for all $t_k \geq t_0$ (c_1 and c_2 are constants).

In (5, 244-251) Jazwinski also discusses error sensitivity. In this latter regard, he proved the following:

Theorem. If $P_0 \leq P_0^c$, $Q_k \leq Q_k^c$, and $R_k \leq R_k^c$ for all k, then $P_k \leq P_k^c$ and $\overline{P}_k \leq \overline{P}_k^c$ for all k.

The superscript, c, refers to the numerical values employed or computed in the filtering algorithm. The non-superscripted quantities refer to the actual or true values of the covariances. Note that the true values are seldom known in practice. Hence, in light of this theorem we can enjoy the confidence that if conservative values are selected for the covariances, then the true error covariance is bounded at any time by the computed error covariance.

1.4 The Problem of Modeling Errors

Up to this point in the discussion of Kalman filtering it has been assumed that the system dynamics are known to within a degree of uncertainty represented by the statistics associated with the state noise, ω. If the uncertainties in the model dynamics are purely random with accurately known finite, bounded statistics, and more realistically if the uncertainties are small with respect to the state values, then generally the assumed dynamic modeling will yield good results. On the other hand, if there are

modeling errors which contain unbounded time correlated components, or dynamic biases as they are usually called, the estimation results based on such erroneous models become worthless. In such cases, the system is not uniformly completely controllable. In particular, the model errors manifest themselves in the Φ-matrix which affects the predicted values of the state, $\bar{x}$, and error covariance, $\bar{P}$. If the filter employs no process noise covariance, Q, or only a constant Q, then as the estimation process proceeds, the state error covariance usually* decreases and so, therefore, does the gain, K. As the gain becomes smaller, the effects of the measurements in contributing information become less and less. The effects of modeling errors continue to grow, essentially no new information is added, and the state estimate diverges from its correct value. A simple example will be developed which will serve to illustrate these points. It is derived from information given in reference (9).

Consider two spacecraft in orbit about the earth. Assume their orbits are approximately circular and coplanar, and that the vehicles are undergoing a rendezvous maneuver, and hence are separated by a distance of only a few kilometers. Referring to Figure 1.3, let r_k be the geocentric position vector of the target vehicle and r_h be the geocentric position vector of the homing vehicle. Neglecting n-body and aspherical gravitational effects, the equations of motion can be written as

$$\ddot{r}_t = g(r_t) = -\frac{\mu}{|r_t|^3}\, r_t$$

$$\ddot{r}_h = g(r_h) = -\frac{\mu}{|r_h|^3}\, r_h \qquad (1.4.1)$$

---

* Recall that $||\Phi_{k,0}||^2$ in *(1.3.16)* must not increase faster than $||M_{k,1}^{-1}||$ decreases. A similar characteristic is required in the system with process noise.

where $(\dot{}) = d(\ )/dt$, μ = gravitational parameter of the earth, and $|r|$ is the magnitude of r.

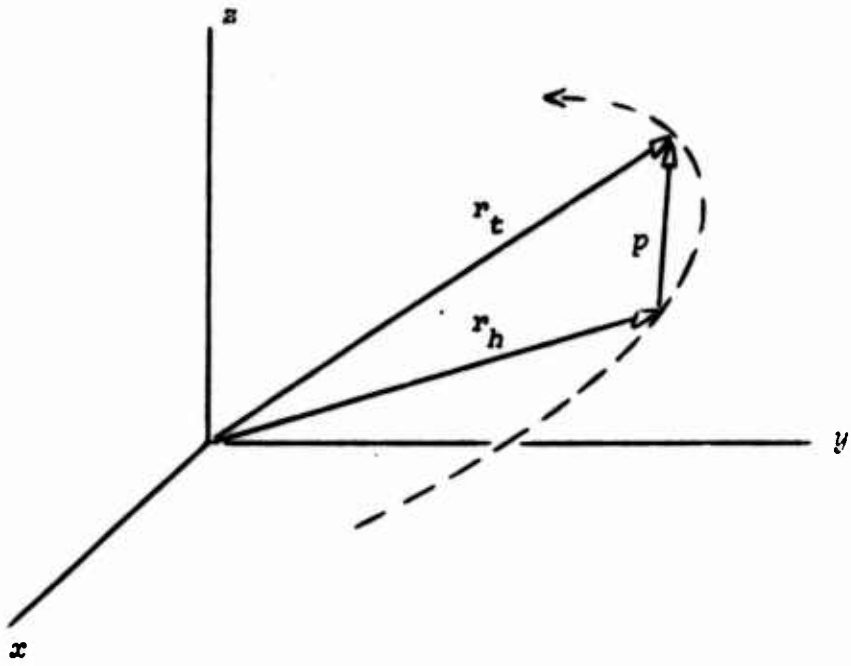


Figure 1.3 Vehicle Position Vectors

Defining the range vector,

$$p = r_t - r_h \tag{1.4.2}$$

and noting that $|p| << |r_h|$ or $|r_t|$, one can expand about the target vehicle position vector in a Taylor series to obtain

$$g(r_h) = g(r_t) + \frac{\partial g(r_t)}{\partial r_t}(r_h - r_t) + \varepsilon \tag{1.4.3}$$

Neglecting ε, which represents terms of second and higher order in p, and using (1.4.1) and (1.4.2) leads to the following relation:

$$\ddot{p} = \frac{\partial g(r_t)}{\partial r_t}\, p \tag{1.4.4}$$

Note that neglecting ε, as well as the higher order gravitational effects

has introduced errors in the dynamic model. However, for the time being assume that *(1.4.4)* represents the true model, that the orbits are coplanar, and that the target vehicle orbit is circular with the homing vehicle orbit nearly so. Further, for purposes of developing a simple navigation scheme the possibly questionable assumption will be made that the homing vehicle is closing at a constant rate, $\dot{\rho}$, with respect to the target vehicle. Note that $\dot{\rho}$ is the magnitude of the time rate of change of the range vector, ρ, i.e.,

$$\dot{\rho} = \frac{\rho \cdot \dot{\rho}}{\rho}$$

where $\dot{\rho} = d\rho/dt$. Throughout the rendezvous maneuver, it is necessary to have knowledge of $\dot{\rho}$ and ρ; however, for simplicity in the illustration, only $\dot{\rho}$ will be considered.

To obtain information about $\dot{\rho}$ a Doppler radar system is used to measure the closure rate. Associated with this range-rate radar are uncertainties assumed to manifest themselves as bounded, purely random errors, v_k, in the range-rate measurements. Assume the v_k have the statistics

$$E\{v_k\} = 0, \; E\{v_k \, v_j^T\} = R \, \delta_{jk}, \; R = constant$$

Making the following identifications with equations *(1.2.1)* and *(1.2.2)*,

$$P_0 = P(0), \; x_k \equiv \dot{\rho}_k, \; H_k \equiv 1, \; y_k = \dot{\rho}_k + v_k$$

and assuming no state noise *(Q = 0)*, then the Kalman estimation equations for Algorithm I are

$$\begin{aligned}
\bar{\dot{\rho}}_k &= \hat{\dot{\rho}}_{k-1} \\
\bar{P}_k &= P_{k-1} \\
K_k &= \bar{P}_k [\bar{P}_k + R]^{-1} \\
\hat{\dot{\rho}}_k &= \bar{\dot{\rho}}_k + K_k [y_k - \bar{\dot{\rho}}_k] \\
P_k &= [1 - K_k] \, \bar{P}_k
\end{aligned} \qquad (1.4.6)$$

Recall that y_k is the measured range-rate at t_k, and that $v_k = y_k - \dot{\hat{\rho}}_k$. Explicit expressions may be obtained for the above quantities in terms of $\dot{\hat{\rho}}_0$, P_0, and y_k. In particular, the recurrence for P_k leads to

$$P = \frac{P_0 R}{(mP_0 + R)} \tag{1.4.7}$$

and

$$K_m = \frac{P_0}{(mP_0 + R)} \tag{1.4.8}$$

for m observations. Then for $\dot{\hat{\rho}}_m$,

$$\dot{\hat{\rho}}_m = \dot{\hat{\rho}}_0 + \sum_{k=1}^{m} \frac{v_k}{(k + R/P_0)} \tag{1.4.9}$$

From equations *(1.4.7)-(1.4.9)*, it follows that as m gets larger, the estimate for $\dot{\hat{\rho}}_m$ becomes insensitive to the observation residual since the variance and hence the gain are becoming smaller. The implication is that eventually a point is reached where K is so small that taking further observations adds essentially no new information about $\dot{\rho}$. This is an acceptable (and desirable) situation if no modeling errors are present, that is if the assumption about a constant closure rate is valid.

To investigate this assumption further, consider equation *(1.4.4)* in terms of the relative derivatives of p.

$$\ddot{p} = \overset{\infty}{p} + 2\,\omega \times \dot{p} + \dot{\omega} \times p + \omega \times (\omega \times p) \tag{1.4.10}$$

where ω is the angular velocity of the line-of-sight, p, and $\dot{(\ )}$ is the time derivative relative to the target vehicle. Using equation *(1.4.4)* one can find

$$\overset{\infty}{p} = \frac{\partial g(r_t)}{\partial r_t} p - 2\,\omega \times \dot{p} - \dot{\omega} \times p - \omega \times (\omega \times p) \tag{1.4.11}$$

From the first of *(1.4.1)*,

$$\frac{\partial g(r_t)}{\partial r_t} = -\frac{\mu}{|r_t|^3}\left(I - \frac{3}{|r_t|^2}\, r_t\, r_t^T\right) \tag{1.4.12}$$

So,

$$\ddot{p} = -\frac{\mu}{|r_t|^3}\left(I - 3\,\frac{r_t\, r_t^T}{|r_t|^2}\right) p \tag{1.4.13}$$

and

$$\overset{\infty}{p} = -\frac{\mu}{|r_t|^3}\left(I - 3\,\frac{r_t\, r_t^T}{|r_t|^2}\right) p - 2\,\omega \times \dot{p} - \dot{\omega} \times p - \omega \times (\omega \times p) \tag{1.4.14}$$

If the rendezvous maneuver is restricted such that p maintains a constant angle with respect to r_t, then ω is constant. Further, because the target vehicle is in circular orbit,

$$\omega = \sqrt{\mu/|r_t|^3}\;\; \hat{k} \tag{1.4.15}$$

where $\hat{k}$ is a unit vector perpendicular to the orbit plane. In view of this restriction,

$$\overset{\infty}{p} = -\frac{\mu}{|r_t|^3}\left(I - 3\,\frac{r_t\, r_t^T}{|r_t|^2}\right) p - 2\,\omega \times \dot{p} - \omega \times (\omega \times p) \tag{1.4.16}$$

Now, take the scalar product of *(1.4.16)* with a unit vector in the p-direction to get

$$\frac{\overset{\infty}{p} \cdot p}{\rho} = \ddot{\rho} = -\frac{\mu}{|r_t|^3}\,\frac{p^T}{\rho}\left(I - 3\,\frac{r_t\, r_t^T}{|r_t|^2}\right) p - 2\,\frac{p^T}{\rho}\,\omega \times \dot{p} - \frac{p^T}{\rho}\,\omega \times (\omega \times p)$$

$$= -\frac{\mu}{|r_t|^3}\,\rho + \frac{3\mu(p^T\, r_t)^2}{\rho |r_t|^5} + \omega^2\,\rho \tag{1.4.17}$$

since $p^T(\omega \times \dot{p}) = p \cdot (\omega \times \dot{p}) = \omega \cdot (\dot{p} \times p) = 0$, and $p^T[\omega \times (\omega \times p)] = (\omega \cdot p)^2 - \omega^2\,\rho^2 = -\omega^2\,\rho^2$. But $\omega^2 = \mu/|r_t|^3$; hence,

$$\ddot{\rho} = \frac{3\mu}{\rho|r_t|^5} (p^T r_t)^2 \qquad (1.4.18)$$

Next,

$$p = \rho \begin{bmatrix} \sin(\omega t + \varepsilon) \\ -\cos(\omega t + \varepsilon) \\ 0 \end{bmatrix}, \quad r_t = |r_t| \begin{bmatrix} \cos \omega t \\ \sin \omega t \\ 0 \end{bmatrix}$$

where ε is a small, constant angle representing the deviation of p from the local horizontal at the target vehicle (if $\varepsilon = 0$, p and r_t are perpendicular). Then, since $p^T r_t$ is the scalar product of p and r_t,

$$\begin{aligned} p^T r_t &= \rho|r_t| \left[\sin(\omega t + \varepsilon) \cos \omega t - \cos(\omega t + \varepsilon) \sin \omega t\right] \\ &= \rho|r_t| \left[\sin \omega t \cos \omega t \cos \varepsilon + \cos^2 \omega t \sin \varepsilon - \right. \\ &\qquad \left. \cos \omega t \sin \omega t \cos \varepsilon + \sin^2 \omega t \sin \varepsilon\right] \\ &= \rho|r_t| \sin \varepsilon \end{aligned} \qquad (1.4.19)$$

Finally, through first order terms, and for closure along the line of sight at a constant angle, ε, from the local horizontal,

$$\ddot{\rho} = \frac{3\mu}{\rho|r_t|^5} \rho^2 |r_t|^2 \sin^2 \varepsilon = \frac{3\mu}{|r_t|^3} \rho \sin^2 \varepsilon$$

or

$$\ddot{\rho} - \gamma^2 \rho = 0, \quad \gamma^2 = \frac{3\mu}{|r_t|^3} \sin^2 \varepsilon, \text{ a constant} \qquad (1.4.20)$$

Thus, only in the special case where $\varepsilon = 0$ is the assumption of a constant closure rate valid ($\ddot{\rho} = 0$ to first order). The solution to (1.4.20) is found to be

$$\rho = \rho_0 \cosh \gamma t + \frac{\dot{\rho}_0}{\gamma} \sinh \gamma t \qquad (1.4.21)$$

So,

$$\dot{\rho} = \rho_0 \gamma \sinh \gamma t + \dot{\rho}_0 \cosh \gamma t \qquad (1.4.22)$$

and it is seen that there is an unbounded growth in $\dot{\rho}$. A non-dimensional form of *(1.4.22)* is plotted in Figure 1.4.

Returning to the estimation problem, using the original model (constant $\dot{\rho}$), the observation-state relation becomes

$$y_k = \dot{\rho}_k + v_k = \rho_0 \gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k + v_k$$

and

$$y_k - \bar{\dot{\rho}}_k = \rho_0 \gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k - \bar{\dot{\rho}}_k + v_k \qquad (1.4.23)$$

So the estimate becomes

$$\hat{\dot{\rho}}_k = \bar{\dot{\rho}}_k + \frac{1}{(k + R/P_0)} (\rho_0 \gamma \sinh \gamma t_k + \rho_0 \cosh \gamma t_k - \bar{\dot{\rho}}_k + v_k)$$

$$= \frac{(k-1) + R/P_0}{k + R/P_0} \bar{\dot{\rho}}_k + \frac{v_k}{(k + R/P_0)} +$$

$$\frac{1}{k + R/P_0} (\rho_0 \gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k) \qquad (1.4.24)$$

Forming the error, $\tilde{\dot{\rho}}_k$, by subtracting *(1.4.24)* from *(1.4.22)* gives

$$\tilde{\dot{\rho}}_k = \left[\frac{(k-1) + R/P_0}{k + R/P_0}\right] (\rho_0 \gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k - \bar{\dot{\rho}}_k) - \frac{v_k}{(k + R/P_0)} \qquad (1.4.25)$$

From *(1.4.24)* one can see that as k increases, the additive effects of the last term become less (since $\gamma << 1$), and $\hat{\dot{\rho}}_k$ tends to a relatively constant value. In *(1.4.25)*, the term in square brackets tends to *1*, and since $\bar{\dot{\rho}}_k$ is approximately constant for large k, the hyperbolic trigonometric terms eventually dominate, driving the error to intolerably large values. As before, the computed variance given by *(1.4.7)* decreases to zero.

Consideration of state noise to represent dynamic model uncertainties has been shown to have the effect of keeping the error covariance from

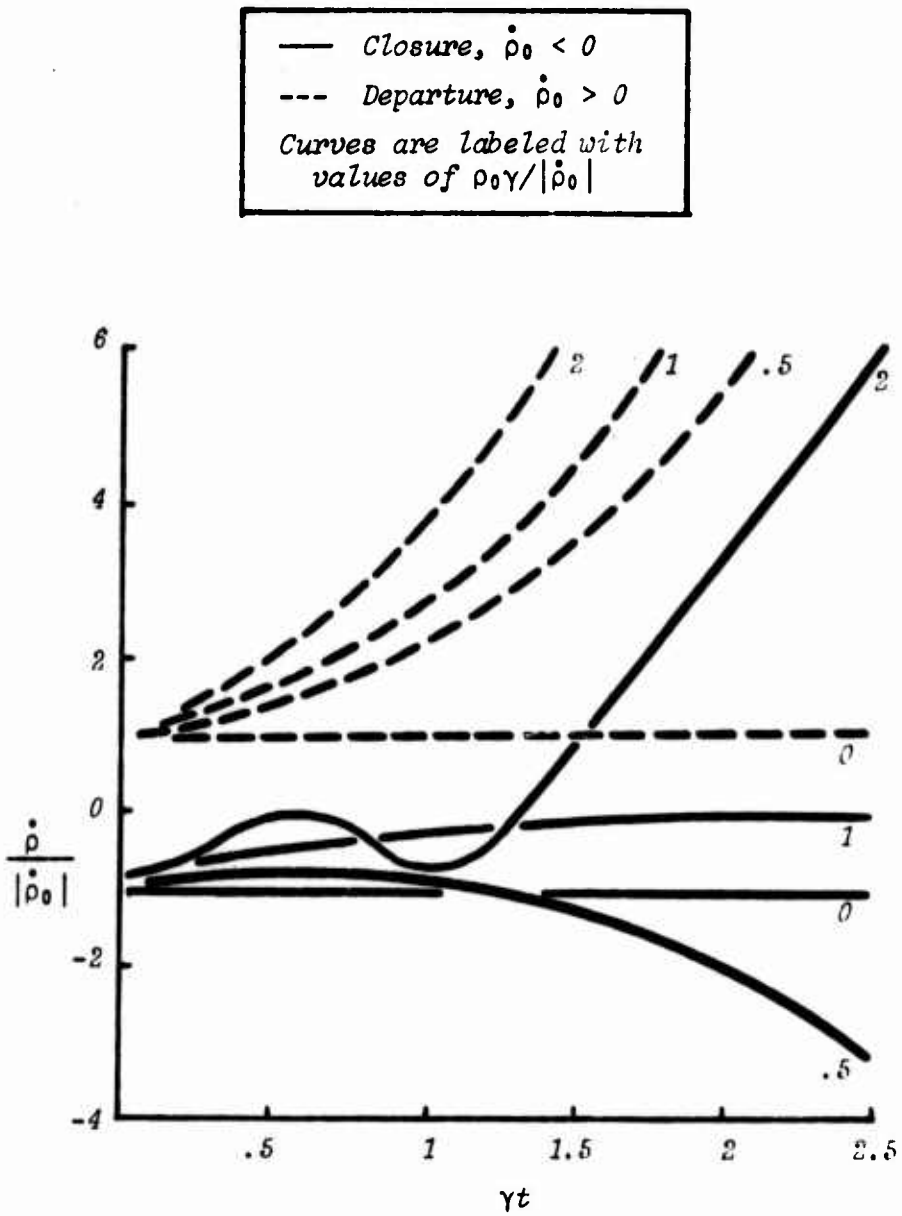


Figure 1.4 Non-Dimensionalized Range Rate

vanishing, and likewise, the gain. For the rendezvous example, let Γ in equation *(1.2.1)* be the identity matrix. Further, let the process generating w_k be stationary so that the variance, $Q_k = Q$, is a positive constant. Then,

$$\bar{P}_k = P_{k-1} + Q \qquad (1.4.28)$$

For $k = 1$,

$$K_1 = \bar{P}_1/[\bar{P}_1 + R] = \frac{P_0 + Q}{P_0 + Q + R}$$

Assuming the initial estimate error is zero, the actual *a priori* variance, P_0, is simply R. Then the simplification may be made that

$$K_1 = \frac{R + Q}{2R + Q} = \frac{1 + f}{2 + f}, \quad f \equiv \frac{Q}{R}$$

and

$$P_1 = \frac{1 + f}{2 + f} R$$

For $k = 2$,

$$\bar{P}_2 = R\left[\frac{1 + f}{2 + f}\right] + Q = R\left[\frac{1 + f}{2 + f} + f\right] = R\left[\frac{1 + 3f + f^2}{2 + f}\right]$$

$$K_2 = \frac{\bar{P}_2}{\bar{P}_2 + R} = \frac{(1 + 3f + f^2)/(2 + f)}{(1 + 3f + f^2)/(2 + f) + 1} = \frac{1 + 3f + f^2}{3 + 4f + f^2}$$

$$P_2 = \left[\frac{1 + 3f + f^2}{3 + 4f + f^2}\right] R$$

For $k = 3$,

$$\bar{P}_3 = \left[\frac{(1 + 3f + f^2)}{(3 + 4f + f^2)} + f\right] R$$

$$= \left[\frac{1 + 6f + 5f^2 + f^3}{3 + 4f + f^2}\right] R$$

$$K_3 = \left[\frac{1 + 6f + 5f^2 + f^3}{1 + 6f + 5f^2 + f^3 + 3 + 4f + f^2}\right] = \frac{1 + 6f + 5f^2 + f^3}{4 + 10f + 6f^2 + f^3}$$

$$P_3 = \left[\frac{1 + 6f + 5f^2 + f^3}{4 + 10f + 6f^2 + f^3}\right] R$$

For $k = 4$,

$$\bar{P}_4 = \left[\frac{1 + 10f + 15f^2 + 7f^3 + f^4}{4 + 10f + 6f^2 + f^3}\right] R$$

$$K_4 = \left[\frac{1 + 10f + 15f^2 + 7f^3 + f^4}{5 + 20f + 21f^2 + 8f^3 + f^4}\right]$$

$$P_4 = K_4 R$$

Finally, for $k = 5$,

$$\bar{P}_5 = \left[\frac{1 + 15f + 35f^2 + 28f^3 + 9f^4 + f^5}{5 + 20f + 21f^2 + 8f^3 + f^4}\right]$$

$$K_5 = \left[\frac{1 + 15f + 35f^2 + 28f^3 + 9f^4 + f^5}{6 + 35f + 56f^2 + 36f^3 + 10f^4 + f^5}\right]$$

$$P_5 = K_5 R$$

Examining these values leads, by induction, to the relationship

$$K_m = \frac{\sum_{k=0}^{m} \binom{m+k}{m-k} f^k}{\sum_{k=0}^{m} \binom{m+k+1}{m-k} f^k} \qquad (1.4.27)$$

where $\binom{a}{b} = \frac{a!}{(a-b)!\, b!}$

the binomial coefficients, and $f = Q/R$. P_m and $\bar{P}_m$ are given as before.

It is important to investigate the limit of K_k as $k \to \infty$. Returning to the recursive relations

$$K_k = \frac{\bar{P}_k}{\bar{P}_k + R}$$

$$P_k = [1 - K_k]\, \bar{P}_k$$

$$\bar{P}_{k+1} = P_k + Q$$

then

$$\bar{P}_{k+1} = \left[1 - \frac{\bar{P}}{\bar{P}_k + R}\right] \bar{P}_k + Q$$

$$= \frac{\bar{P}_k R}{\bar{P}_k + R} + Q \qquad (1.4.28)$$

Now, as $k \to \infty$,

$$\bar{P}_\infty = \frac{\bar{P}_\infty R}{\bar{P}_\infty + R} + Q \qquad (1.4.29)$$

where the notation $\bar{P}_\infty$ means $\lim_{k\to\infty} \bar{P}_k$.

Solving for $\bar{P}_\infty$,

$$\bar{P}_\infty = \frac{Q \pm \sqrt{Q^2 + 4RQ}}{2} \qquad (1.4.30)$$

where the + sign is taken since it is required that $P > 0$. However,

$$K_\infty = \frac{\bar{P}_\infty}{\bar{P}_\infty + R} = \frac{(Q + \sqrt{Q^2 + 4QR})/2}{(Q + \sqrt{Q^2 + 4QR})/2 + R}$$

$$= \frac{1 + \sqrt{1 + 4/f}}{1 + \sqrt{1 + 4/f} + 2/f} \qquad (1.4.31)$$

and therefore,

$$P_\infty = \left[\frac{1 + \sqrt{1 + 4/f}}{1 + \sqrt{1 + 4/f} + 2/f}\right] R \qquad (1.4.32)$$

Noting that $K_\infty < 1$, and $P_\infty < R$, one may write

$$K_\infty < K_k < 1$$

$$P_\infty < P_k < R \qquad (1.4.33)$$

The fact that P and K do not vanish due to the presence of Q does not in general prevent divergence of the estimate, when unbounded model errors are present, but simply delays it. Forming the estimate for very large k , one obtains approximately,

$$\hat{\dot{\rho}}_k \simeq [1 - K_\infty]\, \bar{\dot{\rho}}_k + K_\infty[\rho_0\gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k + v_k],$$

$$\simeq \left[\frac{2/f}{1 + \sqrt{1 + 4/f} + 2/f}\right] \bar{\dot{\rho}}_k +$$

$$\left[\frac{1 + \sqrt{1 + 4/f}}{1 + \sqrt{1 + 4/f} + 2/f}\right] (\rho_0\gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k + v_k) \qquad (1.4.34)$$

The corresponding error is approximately

$$\tilde{\dot{\rho}}_k \simeq \left[\frac{2/f}{1 + \sqrt{1 + 4/f} + 2/f}\right] (\rho_0\gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k - \bar{\dot{\rho}}_k)$$

$$- \left[\frac{1 + \sqrt{1 + 4/f}}{1 + \sqrt{1 + 4/f} + 2/f}\right] v_k \qquad (1.4.35)$$

Thus the non-zero steady-state gain, K_∞, continues to provide information to improve the estimate of $\dot{\rho}$. Note that as Q, and hence f, becomes very large, the error asymptotically approaches the error due to the observation noise. However, after a sufficient amount of time, the estimate still diverges owing to the unboundedness of the model error.

The question remains: "How does one compensate for modeling errors so as to allow a workable estimation algorithm which provides accurate and useful estimates?"

1.5 Literature Survey

The effects of dynamic model errors have been examined by a number of investigators, among them Heffes (10), Schlee, *et al.* (11), Price (12),

Huddle and Wismer (13), and Neal (14). Most of these investigations are concerned primarily with performance degradation of the estimation process (which is the Kalman sequential estimator in all cases), and offer equations which provide a measure of the degradation. These equations are generally based upon the error covariance and are structured in such a way that model errors may be seen through their effects on the error covariance.

As mentioned in Section 1.1, model error compensation techniques may be classed loosely as adaptive or non-adaptive. Techniques from both camps may be further classed according to their basic approach to the problem. On the one hand are techniques which simply alter the actual estimation or filtering equations. These approaches generally attempt to maintain the error covariance, and hence the gain, at a level which will continue to provide corrective information for the state estimate. On the other hand are found methods which attempt to improve knowledge of the dynamic model. The former approaches usually aim at computational simplicity, but pay the price by compromising optimality. The latter methods are usually more in keeping with the optimality properties but generally are computationally more demanding.

The technique of representing modeling errors as white noise is an easily implemented approach which is perhaps the most conservative. Numerically, it amounts to merely increasing $\bar{P}_k$ by adding the Q_{k-1} matrix to $\Phi P_{k-1} \Phi^T$. As pointed out in the last section, this merely delays divergence if the errors are unbounded, but can be effective for small, bounded errors, i.e., if the system is uniformly completely controllable. The consideration of state noise is employed so often that it has become part of the standard linear, unbiased, minimum variance estimator (5, 194-209), (15). However, in the face of incomplete knowledge of the dynamic model, Q_{k-1} must be guessed.

Rather than arbitrarily select a complete matrix of values to be added to the error covariance, an effective and simple method is simply to scale the error covariance by a scalar. This has been done recently by Tarn and Zaborszky (16). A scalar, $s > 1$, is selected and used to scale P_k after each covariance update. In the application to an inertially navigated glide vehicle, it was concluded that $1 < s < 1.5$, and that satisfactory results were obtained for $s = 1.2$.

Fagin (17) introduced the concept of exponential age-weighting of the observations. The effect is to downgrade the value of old observations so that the most current information dominates the estimate. Fagin's development is based upon a recursive least squares derivation, where

$$J_k = \tfrac{1}{2}[\tilde{x}_k^T \Phi^T(o, k) P_o^{-1} \Phi(o, k) \tilde{x}_k] + \tfrac{1}{2} \sum_{i=1}^{k} [y_i - H_i \Phi(i, k) x_k]^T \bar{R}_i^{-1} [y_i - H_i \Phi(i, k) x_k] \qquad (1.5.1)$$

is minimized with respect to x_k. In the standard least squares approach, $\bar{R}^{-1} = R^{-1}$. However, Fagin modifies R_i as

$$\bar{R}_i = exp\ [(t_k - t_i)/\tau]\ R_i,\ i \geq k \qquad (1.5.2)$$

where τ is an arbitrarily chosen time constant. The resulting estimation equations have a form different from those of Tarn and Zaborszky; however, they are equivalent (this will be shown in Chapter 2 along with certain requirements on the scaling factor).

Miller (18) examines the behavior of the Kalman filter for continuous and discrete time invariant systems with exponentially age-weighted observations. He derives equations for the filter eigenvalues to provide useful guides for choosing the aging time constant, τ.

Schmidt (19), (20) offers two unique methods for overcoming model errors. One utilizes a scalar parameter b to scale the optimum gain, K, with an attendant increase in the updated covariance matrix. A second technique adds a judiciously selected term to the Kalman gain with the corresponding effect of an additive term to the updated covariance. In both methods the values of the parameters involved are related to certain configurations of the Kalman estimation equations. Both techniques effectively over-weight the more recent data.

Limited memory filtering is a useful technique for counteracting model errors. Here, only a limited batch of observations are employed. However, the "batch" is updated in the sense that old observations are effectively discarded as new observations are added. Hence the dynamic model is required to be commensurate with the data only over a short time interval. While a number of limited memory filters have been developed, one of the most efficient and easiest to implement is due to Jazwinski (21).

Adaptive techniques form the majority of model error compensation methods. Here attempts are made to estimate a parameter vector, a, whose elements are unknowns in the model. These may include elements in Φ, H, Q, or R. Mehra (22) places adaptive methods into four categories.

1. Bayesian
2. Maximum likelihood
3. Correlation
4. Covariance matching

Bayesian methods involve determining the *a posteriori* probability density function, $p(x_k, a|Y_k)$. Employing Bayes' Theorem,

$$p(x_k, a|Y_k) = p(x_k|a, Y_k)\, p(a|Y_k) = p(x_k|a, Y_k)\, \frac{p(Y_k|a)\, p(a)}{p(Y_k)} \qquad (1.5.3)$$

Since

$$p(Y_k) = \int_A p(Y_k|a)\, p(a)\, da,$$

then

$$p(x_k, a|Y_k) = p(x_k|a, Y_k)\, \frac{p(Y_k|a)\, p(a)}{\int_A p(Y_k|a)\, p(a)\, da} \qquad (1.5.4)$$

where A is the set of all a. Usually a recursive formulation is desirable. Noting that

$$p(Y_k) = p(y_k|Y_{k-1})\, p(Y_{k-1}) \qquad (1.5.5)$$

substitution of (1.5.5) into (1.5.4) yields

$$p(x_k, a|Y_k) = p(x_k|a, Y_k)\, \frac{p(y_k|Y_{k-1}, a)\, p(a|Y_{k-1})}{\int_A p(y_k|Y_{k-1}, a)\, p(a|Y_{k-1})\, da} \qquad (1.5.6)$$

where $p(Y_{k-1})$ has been divided out of the numerator and denominator. Now the optimal estimate of x_k is the conditional mean (22, XIV.1.1)

$$\hat{x}_k = E\{x_k|Y_k\} = \int_{X_k} x_k\, p(x_k|Y_k)\, dx_k$$

and for the problem at hand,

$$\hat{x}_k = \int_{X_k} \int_A x_k\, p(x_k|a, Y_k)\, dx\, p(a|Y_k)\, da$$

or

$$\hat{x}_k = \int_A \hat{x}_k(a)\, p(a|Y_k)\, da \qquad (1.5.7)$$

Using the quotient in (1.5.4) substituted for $p(a|Y_k)$ completes the formulation of the algorithm.

The obvious difficulties in *(1.5.4)* are the evaluation of the $p(a|Y_k)$ integrals over A. Magill (25) develops the above algorithm approximating the integrals by summations of discrete probability distribution functions. Thus,

$$\hat{x}_k \simeq \sum_{i=1}^{n} \hat{x}_k(a_i)\ P(a_i|Y_k)$$

and

$$P(a_i|Y_k) \simeq \frac{p(y_k|Y_{k-1},\ a_i)\ P(a_i|Y_{k-1})}{\sum_{j=1}^{n} p(y_k|Y_{k-1},\ a_j)\ P(a_j|Y_{k-1})} \qquad (1.5.8)$$

where *(1.5.5)* has been employed. This algorithm can be effective if the dimension of a is not too large. To implement the procedure, the values of the a_i must be precomputed; therefore, large n increases the computational load. Further, the appropriate probability density and distribution functions must be assumed.

Maximum likelihood estimation is based upon maximizing a likelihood function, L, formed as

$$L = \ell n\ p[x_k,\ a|Y_k]$$

with respect to the state, x, and the parameter vector, a. Taking partial derivatives with respect to the appropriate elements produces a set of usually nonlinear algebraic equations. Thus the disadvantage here is that an iterative method is required to solve for the estimates, or else some approximation must be made which yields suboptimal estimates. Further, the density function must be known, *a priori*. Alternately, using the marginal density function, $p(a|Y_k)$, it is sometimes easier to derive an estimator for a.

The basic approach of correlation methods is to relate the auto-correlation function of the observations to the unknown parameters. These methods are normally applicable only to constant coefficient systems and therefore will not be discussed further. The interested reader should see (29) and (30).

Covariance matching techniques attempt to make the observed residuals match their theoretical covariances through the appropriate choice of the unknown parameters. Approximating the theoretical covariance of the residuals by the sample covariance, one has

$$V \simeq \overline{V}_m = \frac{1}{m} \sum_{i=1}^{m} \nu_i \nu_i^T \qquad (1.5.9)$$

where V is the theoretical covariance, $\overline{V}$ is the sample variance, and ν is the vector of observation residuals. The limit, m, is chosen to provide a semblance of smoothing. Equating V and $\overline{V}$,

$$V = H_k \left[\Phi P_{k-1} \Phi^T + \Gamma Q_k \Gamma^T\right] H_k^T + R = \overline{V}_k \qquad (1.5.10)$$

Care must be taken to insure rank is consistent with the number of unknowns. For example, if Q is being estimated, HP must be of equal or greater rank than Q in order to obtain a unique solution for the elements of Q. Alternatively, additional equations may be obtained for different times. For elements of, say, Φ, *(1.5.10)* is nonlinear, hence an iterative method is generally required. In passing, it is mentioned that covariance matching techniques appear to give fair results (24), and seem to be easiest to implement.

The adaptive techniques discussed above have been used extensively to estimate the state noise covariance matrix, as well as other parameters. We have already mentioned Magill's use of the Bayesian approach to develop an

adaptive algorithm for estimating parameters which are allowed to take on only a finite number of values.

Maximum likelihood estimates of the state noise covariance are obtained by Abramson (26), Sage and Husa (27), and Levy (28) to name a few. Abramson provides a complete and detailed development of optimal and sub-optimal methods for simultaneous estimation of the state and of the noise statistics. While his approach offers estimates of the diagonal elements of Q, Sage and Husa (27) have extended the approach to yield estimates of all elements of Q, although proof of convergence is not established. Using the Sage-Husa algorithm, Levy has constructed a reprocessing filter which periodically reprocesses the accumulated data to obtain increasingly improved estimates of Q (and R). In the same work, Levy points out some shortcomings of the Sage-Husa approach, and subsequently produces a corrected iterative algorithm for simultaneous estimation of Q and R.

Mehra (29) uses the innovations correlation approach to provide estimates of Q and R for time invariant systems. The method is limited to cases for which the number of unknown elements of Q is less than $n \times p$, where n is the state vector dimension, and p is the observation vector dimension. In cases where this restriction is violated, the Kalman gain may be estimated directly, although this alternative utilizes an iterative approach. In (30) Mehra applies the observations correlation approach to the identification of time invariant system parameters.

Jazwinski (31) uses a maximum likelihood approach to develop a state noise covariance estimator. The joint probability density function of m residuals is maximized with respect to Q. For the case of one residual, normally distributed, the result is identical with that for covariance matching. One residual is not a meaningful statistical sample; however, using the

sample variance of a larger residual sample, Jazwinski has obtained satisfactory results (24).

An effective approach to adaptively compensating for modeling errors is to consider the state noise as being correlated in time, and modeled by the first order Langevin differential equation,

$$\dot{w}(t) = Cw(t) + u(t) \qquad (1.5.11)$$

where C is a diagonal matrix of constants, and $u(t)$ is a white noise process with statistics,

$$E[u(t)] = 0,\ E[u(t)\ u^T(s)] = U(t)\ \delta(t - s)$$

By extending the state vector, x, to include the elements of w, the problem resumes its original form with the only difference being a larger state vector. The specification of C is generally not an easy matter; however, by also including its elements in the state vector, the value of C may be adaptively estimated. This approach has been used successfully by Ingram (32) in representing the effects of time correlated random accelerations acting on the Apollo spacecraft. Schutz (33) has used the same approach to account for the effect of mascons in the lunar gravitational field. Tapley and Hagar (34) have successfully utilized equation *(1.5.11)* as well as the second order equation,

$$\ddot{w}(t) = Cw(t) + u(t) \qquad (1.5.12)$$

to represent time correlated uncertainties in the thrust acceleration vector of a continuous low thrust, solar-electric spacecraft. The possibility of utilizing higher order models to represent time correlated noise is also indicated in (34).

A particular disadvantage of each of these techniques is that unless the constant matrix, C, has some state noise variance associated with it, it will settle to some constant value since its associated gain will vanish.

As a result, unless the correlated noise represents the model errors exactly, divergence of the estimate will eventually occur (35).

It is obvious that there is great latitude in selecting functional forms to represent auto-correlated model errors, and these range from the use of simple constants by Russell and Curkendall (36) to higher order differential equations (37). An obvious problem is that of selecting the correct functional form or structure. This problem has been touched on briefly by Lainiotis (38), and will be considered further in Chapter 3.

1.6 Outline of the Investigation

As stated in Section 1.1, this study investigates the utility of a variety of model error compensation techniques and provides a comparison of the effectiveness of these methods. The particular techniques investigated are selected from those mentioned in the previous section, along with several extensions and new approaches. The study is limited to implementation using the Kalman sequential estimators (Algorithms I and II) presented in section 1.2, and the error compensation methods studied are general enough to be applicable to time varying as well as time invariant systems.

In Chapter 2 the suboptimal non-adaptive methods of the Schmidt modifications and of age-weighting devices are presented. Also presented in this chapter is Jazwinski's limited memory filter algorithm.

Chapter 3 is concerned with adaptive methods, both optimal and sub-optimal. Jazwinski's adaptive technique for estimating the state noise covariance is presented. Several extensions are given, and the covariance matching technique applied to yield adaptive forms for the Schmidt and age-weighting algorithms presented in Chapter 2. The briefly mentioned problem of structurally adaptive filtering is considered, and results are obtained

for certain restricted forms. A new method is developed for estimating the state noise covariance matrix. The resulting algorithm employs the linear minimum variance sequential estimator to adaptively estimate the state noise covariance elements for the case of linear observations. The same approach is shown to be capable of modifying the error covariance elements directly.

Chapters 4 and 5 apply the previously presented algorithms to two dynamic systems. The scalar rendezvous problem presented in section 1.4 is used to provide tractable, closed-form solutions and to offer insight into the nature of many of the methods. In Chapter 5 a second problem of a more complex but realistic nature is investigated using the most promising approaches indicated in Chapter 4. This concerns estimating the heliocentric trajectory of a continuous low thrust, solar-electric spacecraft, subject to errors in the thrust acceleration vector. A complete description of this problem is given in Chapter 5.

Chapter 6 is a concise summary of the investigation and presents the general conclusions and recommendations for further study.

Chapter 2

NON-ADAPTIVE METHODS

2.1 Age-Weighting of Data

The rationale for age-weighting the observation data to compensate for model errors is this: The Kalman filter uses all the data as information for obtaining an estimate. However, the dynamic model assumed in the process is in error and therefore yields a reasonable approximation to the true motion over only a finite interval of time. Therefore, attempting to make the data consistent with an erroneous model over the total estimation period is unreasonable over a long period. A logical alternative is to downgrade the importance of the older data. Fagin (17) does precisely this by minimizing ϵ_k in equation *(1.5.1)* where $\bar{R}_i = exp[(t_k - t_i)/\tau]\, R_i$, with τ arbitrarily chosen. The results are equations *(2.1.2)* with $s = exp[(t_k - t_i)/\tau]$.

For the derivation here, a slightly different and more direct approach due to Tarn and Zaborszky (16) is given. Here, a general function, s, is used as the weighting factor instead of the special case of an exponential. Proceeding heuristically, it is noted that aging the old data (decreasing its importance) will yield the same results as increasing the importance of the more current data. This may be accomplished simply by scaling the error covariance by s, with $s > 1$. Thus the error covariance and hence the gain is increased so that more importance is attached to the more recent observations. The subo timal covariance, denoted as P^s, is sP with the Kalman equations modified by simply substituting P^s for P. Thus,

$$P^{s}_{k-1} = \varepsilon P_{k-1}$$

$$\bar{P}_k = \Phi P^{s}_{k-1} \Phi^T + \Gamma Q_{k-1} \Gamma^T$$

$$\bar{x}_k = \Phi \hat{x}_{k-1}$$

$$K_k = \bar{P}_k H^T [H \bar{P}_k H^T + R]^{-1}$$

$$\hat{x}_k = \bar{x}_k + K_k [y_k - H \bar{x}_k]$$

$$P_k = (I - K H) \bar{P}_k \qquad (2.1.1)$$

These are the results arrived at by Tarn and Zaborszky. We note in passing that since their results contain both the additive state noise covariance and the ε-factor, there is some redundancy. The presence of the ε-factor simply slows the decrease in the error covariance, and the algorithm is therefore suboptimal. Setting $Q = 0$ in (2.1.1) and using the definition of P^s one easily obtains Fagin's form,

$$\bar{P}_k = \Phi P^{s}_{k-1} \Phi^T$$

$$\bar{x}_k = \Phi \hat{x}_{k-1}$$

$$K_k = \bar{P}_k H^T [H \bar{P}_k H^T + \frac{R}{\varepsilon}]^{-1}$$

$$\hat{x}_k = \bar{x}_k + K_k [y_k - H \bar{x}_k]$$

$$P^{s}_k = \varepsilon (I - KH) \bar{P}_k \qquad (2.1.2)$$

where $\varepsilon = exp\ [(t_k - t_{k-1})/\tau]$.

The effects on the age-weighting algorithm for various values of ε are now investigated. First, if $\varepsilon = 1$, equations (2.1.1) and (2.1.2) are just the Kalman equations of Algorithm I. If ε is very large, the gain approaches a value such that $HK = 1$, in which case only the most recent

Reproduced from best available copy.

observation is considered in forming the estimate. This latter fact is seen by premultiplying *(1.2.8)* by H.

To gain further insight into the algorithm, *(2.1.2)* is used to form the variance between a perfect measurement and a computed measurement (for a single observation), HP^sH^T. For a perfect observation, $y = Hx$; for the computed observation, $\hat{y} = H\hat{x}$. Differencing these and taking the expected value gives

$$\tilde{y} = y - \hat{y} = H(x - \hat{x}) = H\,\tilde{x}$$

$$E[\tilde{y}\,\tilde{y}^T] = H\,P^s\,H^T$$

which represents a measure of the estimation performance for comparison with the Kalman filter. For the case of scalar observations (to which the vector case can always be reduced), *(2.1.2)* is used to obtain

$$H\,P^s\,H^T = s\left[H\,\overline{P}\,H^T - \frac{(H\,\overline{P}\,H^T)^2}{(H\,\overline{P}\,H^T + R/s)}\right] \qquad (2.1.3)$$

Introducing the definition

$$c = R/H\,\overline{P}\,H^T > 0, \qquad (2.1.4)$$

equation *(2.1.3)* becomes

$$H\,P^s\,H^T = \frac{sR}{s + c} \qquad (2.1.5)$$

For the optimal Kalman filter, $s = 1$, and

$$H\,P\,H^T = \frac{R}{1 + c} \qquad (2.1.6)$$

Forming the ration of *(2.1.5)* and *(2.1.6)*,

$$\frac{sR}{s + c} \cdot \frac{1 + c}{R} = \frac{s + sc}{s + c} \geq 1,\ s \geq 1 \qquad (2.1.7)$$

Thus, the measure of performance, given by HP^sH^T for age-weighting, is bounded below by that for the optimal filter. To obtain an upper bound, consider *(2.1.5)* in the limit as $s \to \infty$. Employ L'Hospital's rule to find

$$\lim_{s\to\infty} H P^{s} H^{T} = \lim_{s\to\infty} \frac{sR}{s+c} = R \qquad (2.1.8)$$

The values of $HP_k^s H^T$ are then bounded according to

$$0 < \frac{R}{1+c} \leq H P_k^{s} H^{T} < R, \; s \geq 1 \qquad (2.1.9)$$

for $k > 1$ (after the first observation).

The variance of perfect measurements, HPH^T, forms a useful common denominator for comparison among various algorithms. Letting this quantity be denoted by ε, the corresponding relationship with s may be found. Again considering the case of scalar observations, equation (2.1.3) is manipulated to give the following sequence:

$$H P H^{T} = \varepsilon = s[H \bar{P} H^{T} - H \bar{P} H^{T}(H \bar{P} H^{T} + \frac{R}{s})^{-1} H \bar{P} H^{T}]$$

$$\varepsilon(H \bar{P} H^{T})^{-1} = s[1 - H \bar{P} H^{T}(H \bar{P} H^{T} + \frac{R}{s})^{-1}]$$

$$\varepsilon + \varepsilon(H \bar{P} H^{T})^{-1} \frac{R}{s} = s(H \bar{P} H^{T} + \frac{R}{s}) - s H \bar{P} H^{T}$$

$$s\,\varepsilon + \varepsilon(H \bar{P} H^{T})^{-1} R = R\, s$$

or

$$s = \frac{\varepsilon}{H \bar{P} H^{T}} \left[\frac{1}{1 - \varepsilon/R}\right] \qquad (2.1.10)$$

This result is consistent with (2.1.9) as can be seen by solving (2.1.10) for ε and noting the results as $s \to 1$ and $s \to \infty$. Note that as $\varepsilon \to R$, $s \to \infty$, and the estimate depends more and more upon the latest observations.

Implementation of the standard age-weighting algorithm is straightforward (Figure 2.1), and may be done using either equations (2.1.1) or (2.1.2).

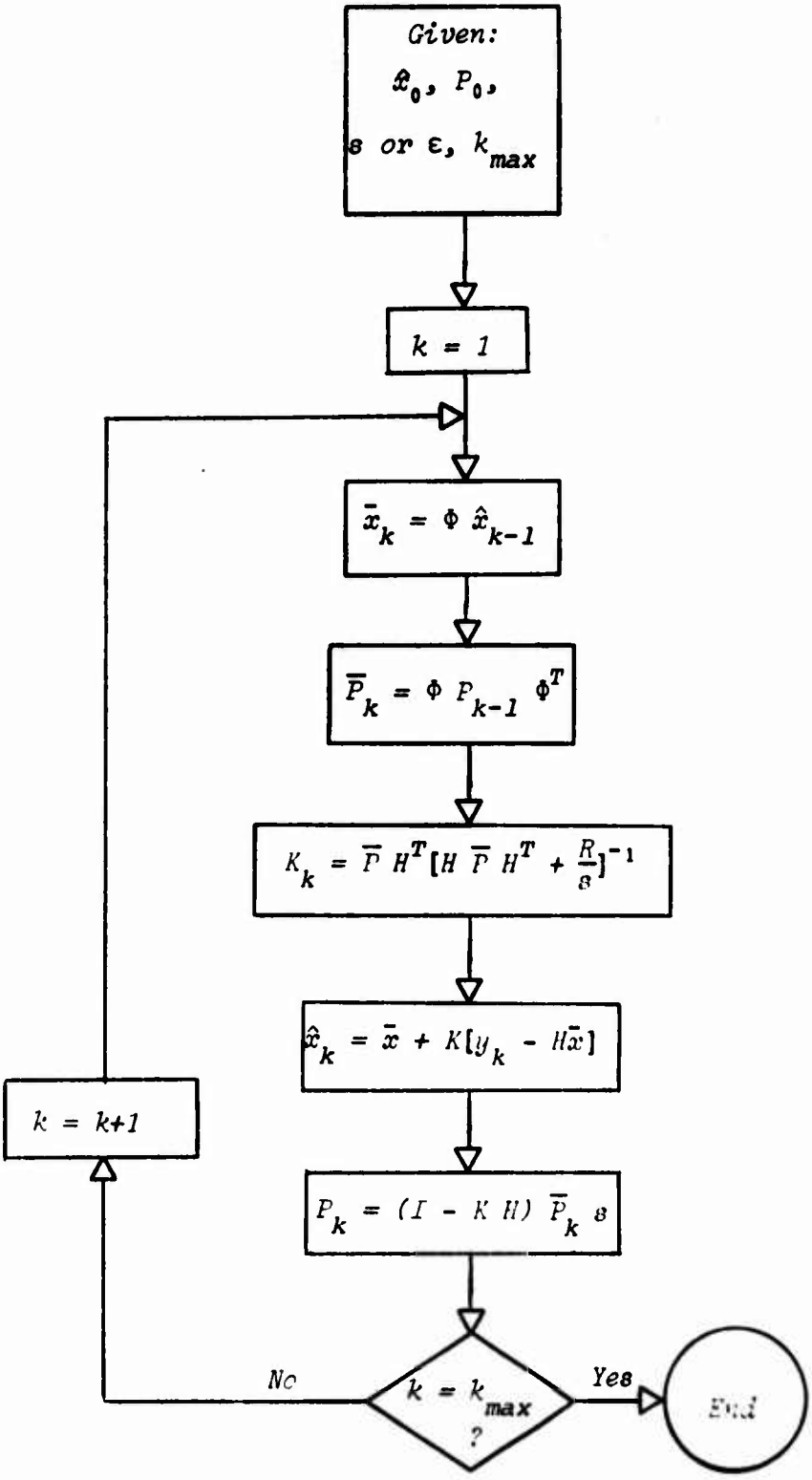


Figure 2.1 Age-Weighted Data Algorithm

2.2 Schmidt Suboptimal Filter: Scaling the Gain

Schmidt (19) considers modifying the gain term directly using a scaling constant, b. The state estimate is then given by

$$\hat{x}_k = \bar{x}_k + b\, K_k[y_k - H\, \bar{x}_k] \tag{2.2.1}$$

and the corresponding error covariance can be found from

$$\tilde{x}_k = x_k - \hat{x}_k = \bar{\tilde{x}}_k - b\, K_k[y_k - H\, x_k] = x_k - b\, K_k[H\, x_k + v_k] \tag{2.2.2}$$

$$P_k = [I - b\, K_k\, H]\, \bar{P}_k [I - b\, K_k\, H]^T + b^2\, K\, R\, K_k^T$$

$$= \bar{P} - b\, K\, H\, \bar{P} - b\, \bar{P}\, H^T\, K^T + b^2\, K(H\, \bar{P}\, H^T + R)\, K^T$$

or

$$P_k = \{I - (2b - b^2)\, \bar{P}_k\, H^T [H\, \bar{P}_k\, H^T + R]^{-1}\, H\}\, \bar{P}_k \tag{2.2.3}$$

Proceeding in a manner similar to that in section 2.1, Schmidt forms

$$H\, P\, H^T = H\, \bar{P}\, H^T - (2b - b^2)\, H\, \bar{P}\, H^T (H\, \bar{P}\, H^T + R)^{-1}\, H\, \bar{P}\, H^T \tag{2.2.4}$$

Considering scalar measurements as before, Schmidt specifies $HP_k H^T$ at some acceptable value, ε, and subsequently solves for b, yielding

$$b = 1 \pm \sqrt{\frac{\varepsilon(H\, \bar{P}\, H^T + R) - R\, H\, \bar{P}\, H^T}{(H\, \bar{P}\, H^T)^2}} \tag{2.2.5}$$

for which real solutions exist when

$$\varepsilon \geq R \left[\frac{H\, \bar{P}\, H^T}{H\, \bar{P}\, H^T + R}\right] = R\, H\, K \tag{2.2.6}$$

and the positive sign in (2.2.5) is taken to insure $b \geq 1$.

The effects of various values of b may be seen readily by examining equation (2.2.3).

$b = 1$ Standard (optimum) Kalman gain and error covariance update equation

$1 < b < 2$ A smaller than optimum term is subtracted from $\overline{P}_k$ in forming the updated covariance, P_k

$b = 2$ $P_k = \overline{P}_k$

$b > 2$ A positive term is added to $\overline{P}_k$

Extending Schmidt's analysis, corresponding values of ε are easily found to be

$$\varepsilon = R\left[\frac{H\,\overline{P}\,H^T}{H\,\overline{P}\,H^T + R}\right] \qquad \text{for } b = 1$$

$$R\left[\frac{H\,\overline{P}\,H^T}{H\,\overline{P}\,H^T + R}\right] < \varepsilon < H\,\overline{P}\,H^T \qquad \text{for } 1 < b < 2$$

$$\varepsilon = H\,\overline{P}\,H^T \qquad \text{for } b = 2$$

$$\varepsilon > H\,\overline{P}\,H^T \qquad \text{for } b > 2$$

Further, values for b and ε may be found which yield $HK \simeq I$, thereby forcing the estimate to depend only upon the more recent observations. In this case,

$$b\,H\,K_k = b\,H\,\overline{P}\,H^T(H\,\overline{P}\,H^T + R)^{-1} \leq 1$$

or

$$b \leq \frac{H\,\overline{P}_k\,H^T + R}{H\,\overline{P}_k\,H^T} \qquad (2.2.7)$$

Using *(2.2.5)* and *(2.2.7)*,

$$\frac{H\,\overline{P}\,H^T + R}{H\,\overline{P}\,H^T} \geq 1 + \frac{\varepsilon(H\,\overline{P}\,H^T + R) - R\,H\,\overline{P}\,H^T}{(H\,\overline{P}\,H^T)^2}$$

one may solve for ε to obtain

$$\varepsilon \leq R \qquad (2.2.8)$$

In view of (2.2.6), (2.2.8) requires that

$$R\left[\frac{H\,\bar{P}_k\,H^T}{R + H\,\bar{P}_k\,H^T}\right] \leq \varepsilon \leq R \qquad (2.2.9)$$

or

$$R\,H\,K \leq \varepsilon \leq R \qquad (2.2.10)$$

For estimates based upon only the current observation, equality must hold in (2.2.8). It may be recalled that this agrees with result obtained for the age-weighting algorithm.

Implementation of this algorithm is straightforward. A slightly different approach involves using the following modified form of (2.2.7):

$$b = \alpha\left[\frac{H\,\bar{P}_k\,H^T + R}{H\,\bar{P}_k\,H^T}\right] \qquad (2.2.11)$$

where $HK \leq \alpha \leq 1$. The lower limit for α follows from the requirement that $b \geq 1$. When applied to the Kalman gain, the following result is obtained:

$$K' \equiv b\,K = \alpha\left[\frac{H\,\bar{P}\,H^T + R}{H\,\bar{P}\,H^T}\right]\frac{\bar{P}\,H^T}{(H\,\bar{P}\,H^T + R)} = \frac{\bar{P}\,H^T}{H\,\bar{P}\,H^T}\,\alpha \qquad (2.2.12)$$

and the covariance update equation (2.2.3) becomes

$$P = \bar{P} - [2\alpha - (1 + \frac{R}{H\,\bar{P}\,H^T})\,\alpha^2]\,\frac{\bar{P}\,H^T\,H\,\bar{P}}{H\,\bar{P}\,H^T} \qquad (2.2.13)$$

Suppose $\alpha = \varepsilon/R$. From the limits on α, the corresponding limits on ε are seen to be those in (2.2.9). Thus, $\varepsilon = R$ results in estimates based upon only the current observation.

The gain scaling technique using a constant value of b can lead to some problems, and this will be shown specifically in Chapter 4. Another alternative method is to determine b as

$$b = 1 + \beta\,\frac{R}{H\,\bar{P}\,H^T}\quad,\quad 0 \leq \beta \leq 1 \qquad (2.2.14)$$

The limits on β correspond to the expected limits for b discussed earlier. In particular, $\beta = 0$ gives $b = 1$ resulting in the optimal Kalman filter; $\beta = 1$ gives the value of b indicated by equality holding in *(2.2.7)*, i.e., all information is obtained from the most recent observation. Employing *(2.2.14)* in the gain and the error covariance update equation *(2.2.3)* gives the following:

$$K' = b\,K$$

$$= \left[1 + \frac{\beta R}{H \bar{P} H^T}\right] \bar{P} H^T (H \bar{P} H^T + R)^{-1}$$

$$= \left[\bar{P} H^T + \beta R \bar{P} H^T / H \bar{P} H^T\right] (H \bar{P} H^T + R)^{-1} \qquad (2.2.16)$$

$$2b - b^2 = 2 + 2\frac{\beta R}{H \bar{P} H^T} - 1 - 2\frac{\beta R}{H \bar{P} H^T} - \frac{\beta^2 R^2}{(H \bar{P} H^T)^2} = 1 - \frac{\beta^2 R^2}{(H \bar{P} H^T)^2} \qquad (2.2.16)$$

$$P = \left\{I - \left[1 - \frac{\beta^2 R^2}{(H \bar{P} H^T)^2}\right] \bar{P} H^T (H \bar{P} H^T + R)^{-1} H\right\} \bar{P} \qquad (2.2.17)$$

Forming HPH^T yields

$$H P H^T = \frac{(H \bar{P} H^T + \beta^2 R)}{H \bar{P} H^T + R} R \qquad (2.2.18)$$

It has been indicated by Schmidt (20) that this method of determining b yields estimation results identical with his additive gain term algorithm discussed in the next section.

The gain scaling technique for constant b or *(2.2.14)* to determine b is diagrammed in Figure 2.2. Figure 2.3 gives the logic flow for the modified gain scaling algorithm using equations *(2.2.12)* - *(2.2.13)*. For this modified approach, it is possible that a value of ε might be chosen which is initially less than the lower bound given by *(2.2.9)*. This is not generally desirable. The philosophy here is to operate with the ordinary Kalman filter

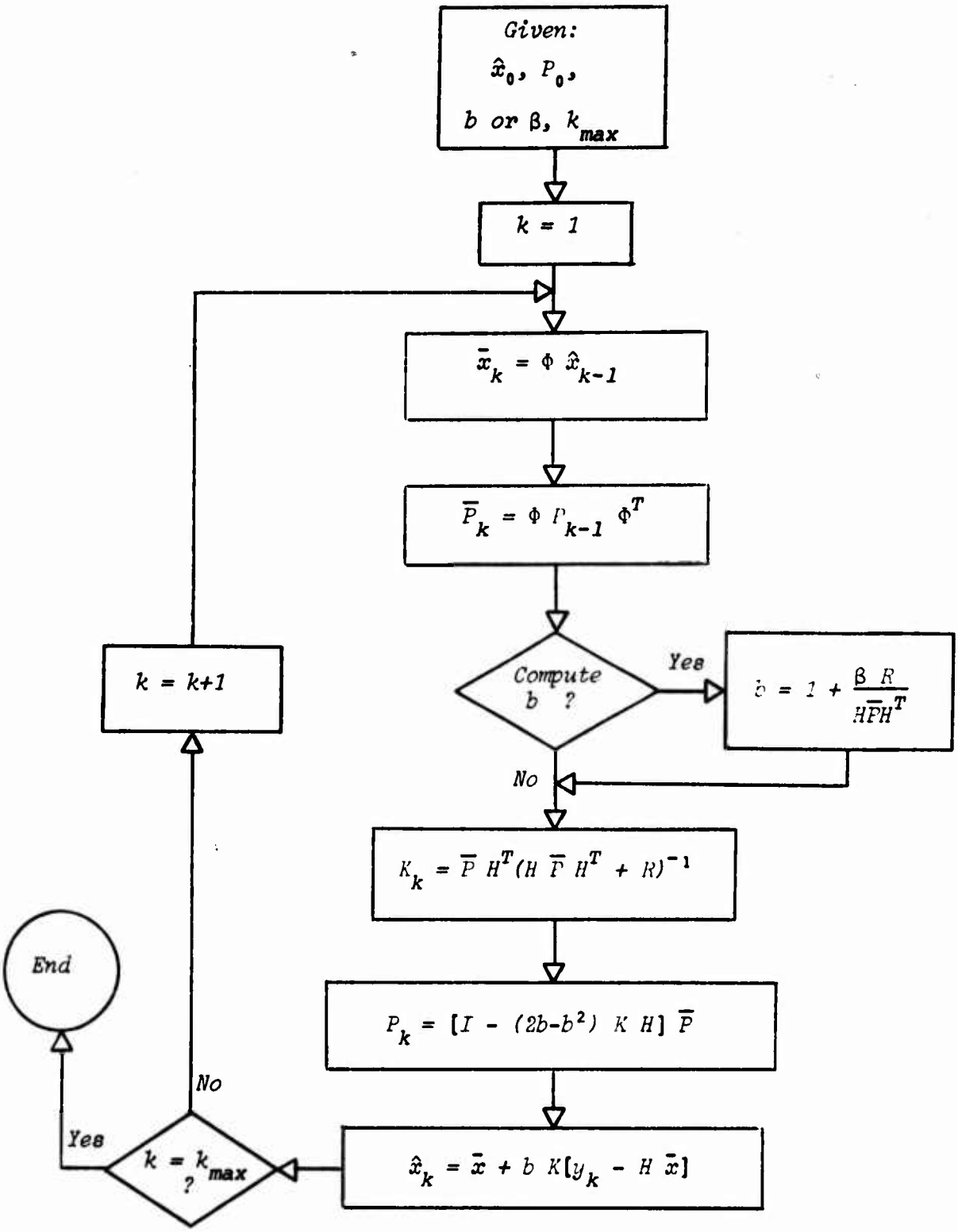


Figure 2.2 Schmidt Suboptimal Filter: Gain Scaling

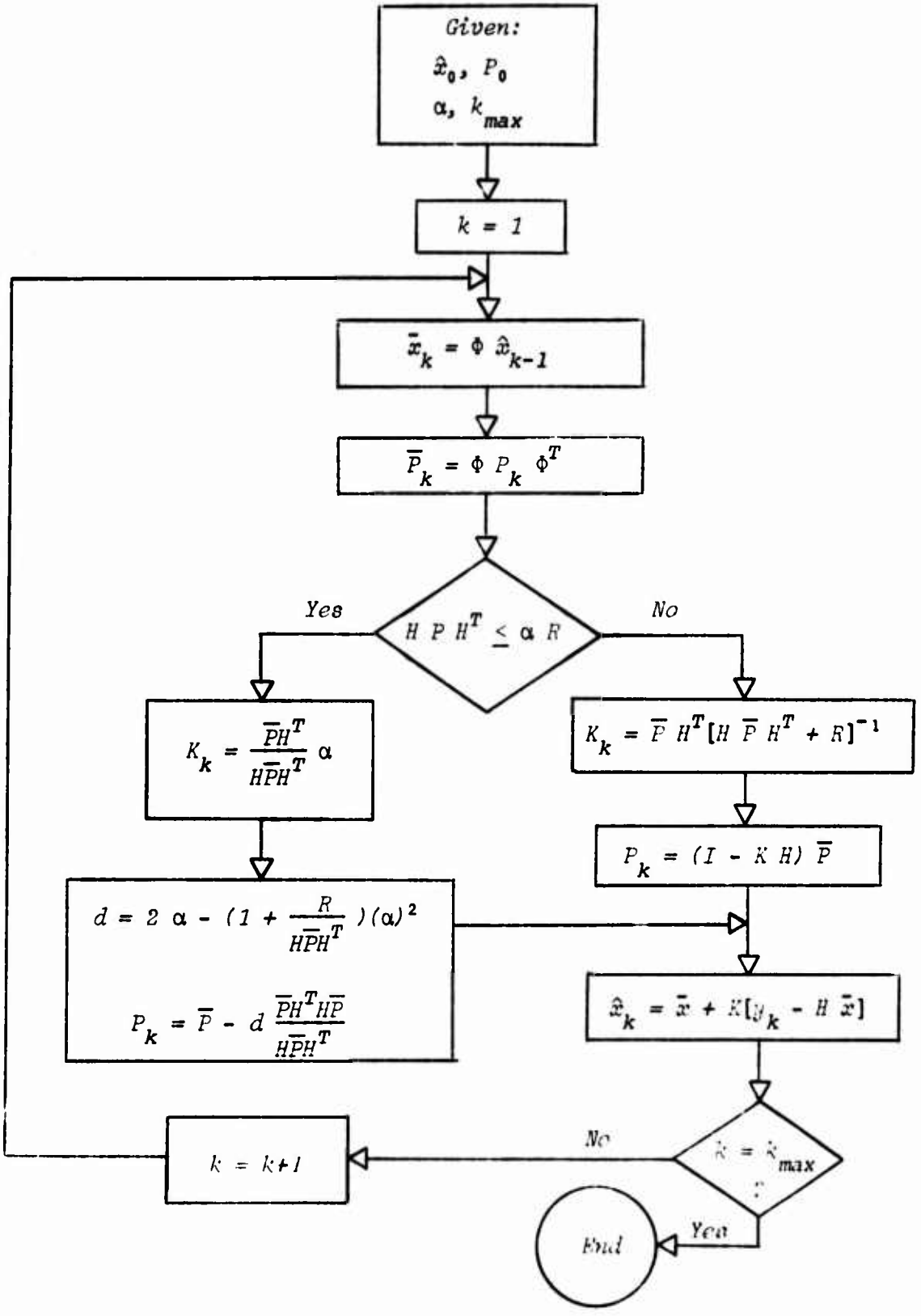


Figure 2.3 Modified Gain Scaling Algorithm

equations, and when $HPH^T = \varepsilon$, switch to the suboptimal modified gain scaling equations.

2.3 Schmidt Suboptimal Filter: Additive Gain Term

A second approach taken by Schmidt (19) is to add a judiciously chosen term, L, to the optimal gain, K. This approach is possible only for scalar observations. The suboptimal term is given as

$$\beta L_k = \beta \frac{R H^T}{(H H^T)\ [H \bar{P}_k H^T + R]} \qquad (2.3.1)$$

so that the actual suboptimal gain is found as

$$\begin{aligned} M_k &= K_k + \beta L_k \\ &= \bar{P}_k H^T [H \bar{P}_k H^T + R]^{-1} + \beta \frac{R H^T}{(H H^T)\ [H \bar{P}_k H^T + R]} \end{aligned} \qquad (2.3.2)$$

If β is chosen as

$$\beta = \beta' \left[\frac{H \bar{P}_k H^T + R}{R} \right]$$

then

$$\beta L = \beta' \frac{H^T}{H H^T}$$

If one considers the estimate of x with no *a priori* information, one obtains

$$\hat{x}_k = \bar{x}_k + H^{\#} [y_k - H \bar{x}_k]$$

where $H^{\#}$ is the pseudo inverse (45, 82-89) of H.

$$H^{\#} = \frac{H^T}{H H^T}$$

Then *(2.3.2)* is simply a linear combination of two gains: the optimal gain for all the observations, and the gain for no *a priori* information, given by the pseudo inverse, $H^{\#}$.

The updated covariance equation is also modified. Using (2.3.2) to form the state estimate,

$$\hat{x}_k = \bar{x}_k + M_k[y_k - H\,\bar{x}_k] \tag{2.3.3}$$

and subsequently the estimate error,

$$\tilde{x}_k = \tilde{\bar{x}}_k - M_k[H\,\tilde{\bar{x}}_k + v_k] \tag{2.3.4}$$

the error covariance is determined as

$$E\{\tilde{x}_k\,\tilde{x}_k^T\} = P_k = (I - M\,H)\,\bar{P}_k(I - M\,H)^T + M\,R\,M^T$$

$$= (I - M\,H)\,\bar{P} - \bar{P}\,H^T\,M^T + M(H\,\bar{P}\,H^T + R)\,M^T$$

But since

$$M(H\,\bar{P}\,H^T + R) = (\bar{P}\,H^T + \beta\,\frac{R\,H^T}{H\,H^T}), \quad P_k = (I - M\,H)\,\bar{P}_k + \beta\,\frac{H^T R}{H\,H^T}\,M^T \tag{2.3.5}$$

Thus the updated covariance matrix is of the same form as the optimal equation with the exception that the suboptimal gain, M, is used, and an additional term is added. An alternative form of (2.3.5) may be found as

$$P_k = (I - K\,H)\,\bar{P}_k + \beta^2\,R\,L\,H/(H\,H^T) \tag{2.3.6}$$

This is the form derived by Schmidt.

Reasonable limits for β are readily established by considering HM, formed using (2.3.2).

$$H\,M = H\,\bar{P}\,H^T(H\,\bar{P}\,H^T + R)^{-1} + \beta\,R(H\,\bar{P}\,H^T + R)^{-1}$$

$$= (H\,\bar{P}\,H^T + \beta\,R)(H\,\bar{P}\,H^T + R)^{-1} \tag{2.3.7}$$

As before, for $HM = 1$, the most recent observation drives the estimate. Hence using (2.3.7)

$$H\,M \leq 1 \Rightarrow \beta \leq 1$$

The lower limit on β is established heuristically by considering the fact that $\beta < 0$ decreases the optimal gain, an effect which counters the idea of

depending more on the recent observations. Thus limits on β are

$$0 \leq \beta \leq 1 \qquad (2.3.8)$$

For $\beta = 0$, $M = K$; for $\beta = 1$ only the current observation is used to determine the estimate.

It is interesting to extend Schmidt's analysis, and find the relationship between b and β. Forming HPH^T from equation (2.3.6),

$$H P H^T = H \bar{P} H^T - (H \bar{P} H^T)^2 (H \bar{P} H^T + R)^{-1} + \beta^2 R^2 (H \bar{P} H^T + R)^{-1}$$

Denoting $\varepsilon = HPH^T$, as in section 2.2, an expression for β may be found.

$$\varepsilon(H \bar{P} H^T + R) = (H \bar{P} H^T)(H \bar{P} H^T + R) - (H \bar{P} H^T)^2 + \beta^2 R^2$$

or

$$\beta = \pm \sqrt{\frac{\varepsilon(H \bar{P} H^T + R) - H \bar{P} H^T R}{R^2}} \qquad (2.3.9)$$

where the positive sign is taken in light of (2.3.8). Now comparing this with the expression for b in section 2.2, b may be rewritten as

$$b = 1 + \beta \frac{R}{H \bar{P} H^T} \qquad (2.3.10)$$

Equation (2.3.10) is identical with (2.2.14), and therefore provides the link between the gain scaling and additive gain term algorithms. It should be noted that this relationship is derived by forming HPH^T, and then developing the equation based on this. By examining equations (2.2.15), (2.3.2), (2.2.17), and (2.3.6) one will see that these gains and covariances differ by the presence or absence of the error covariance matrix in the suboptimal term. The case where it is present, i.e., the suboptimal gain term is given as

$$\frac{\beta R \bar{P} H^T}{(H \bar{P} H^T)(H \bar{P} H^T + R)}$$

is in fact more general. This is true because when the suboptimal gain term is

$$\frac{\beta R H^T}{(H H^T)(H \overline{P} H^T + R)}$$

any zero terms in H prevent contribution to the corresponding optimal Kalman gain term. Usually this does not occur for the former case.

Note that in deriving *(2.3.10)*, when HPH^T is formed for each instance, the differences in the suboptimal covariance terms disappear. Hence these differences do not occur in *(2.3.10)*.

Conditions for β may be established similar to those for b in section 2.2. In particular,

$\beta = 0$	Standard (optimum) Kalman gain and error covariance update equation
$0 < \beta < \frac{H \overline{P} H^T}{R}$	A smaller than optimum term is subtracted from $\overline{P}_k$ in forming the updated covariance, P_k
$\beta = \frac{H \overline{P} H^T}{R}$	$P_k = \overline{P}_k$
$\beta < \frac{H \overline{P} H^T}{R}$	A positive term is added to $\overline{P}_k$

Employing *(2.3.8)* and *(2.3.9)*, limits may be found for ε.

$$\beta = \sqrt{\frac{\varepsilon(H \overline{P} H^T + R) - H \overline{P} H^T R}{R^2}} \leq 1$$

or, using the inequality,

$$\varepsilon \leq R \qquad (2.3.11)$$

which is just the condition established in section 2.2.

As in the previous section, real solutions require the radicand to be positive; i.e.,

$$\leq \frac{K H \bar{P} H^T}{(H \bar{P} H^T + K)}$$

which is just the inequality *(2.2.6)*.

The implementation of this algorithm is given in Figure 2.4.

2.4 Limited Memory Filter

The philosophy behind the limited memory filter is similar to that of the suboptimal techniques presented in the previous sections. However, rather than weight the data in some arbitrary sense, limited memory filters "discard" old observations.

Jazwinski (21) develops a limited memory which is suboptimal only in the sense that it does not take into account the information from all observations in forming the state estimate. However, the filter equations developed are, in fact, optimal over the set of observations considered. In other words, for some subset of the total observations, the limited memory filter provides a linear, minimum variance, unbiased estimate of the state.

Jazwinski's derivation applies to the general, nonlinear case. However, it requires assumption of the probability density function. He also shows that the linear limited memory filter may be derived from least squares considerations. Since concern here is with the linear problem, this approach is sufficient for our purposes. The derivation is duplicated here.

Consider the estimate of the state at time t_k based on observations through time t_m. As defined in Chapter 1, this is denoted as $\hat{x}_{k|m}$. The least squares estimate of x at t_k based on observations through time t_k is derived (5, 206) as follows: Form the performance index, J_k, as

$$J_k = \tfrac{1}{2}(x_m - \hat{x}_{m|m})^T P^{-1}_{m|m}(x_m - \hat{x}_{m|m}) +$$

$$\tfrac{1}{2} \sum_{i=m+1}^{k} (y_i - H_i x_i)^T R_i^{-1}(y_i - H_i x_i) \qquad (2.4.1)$$

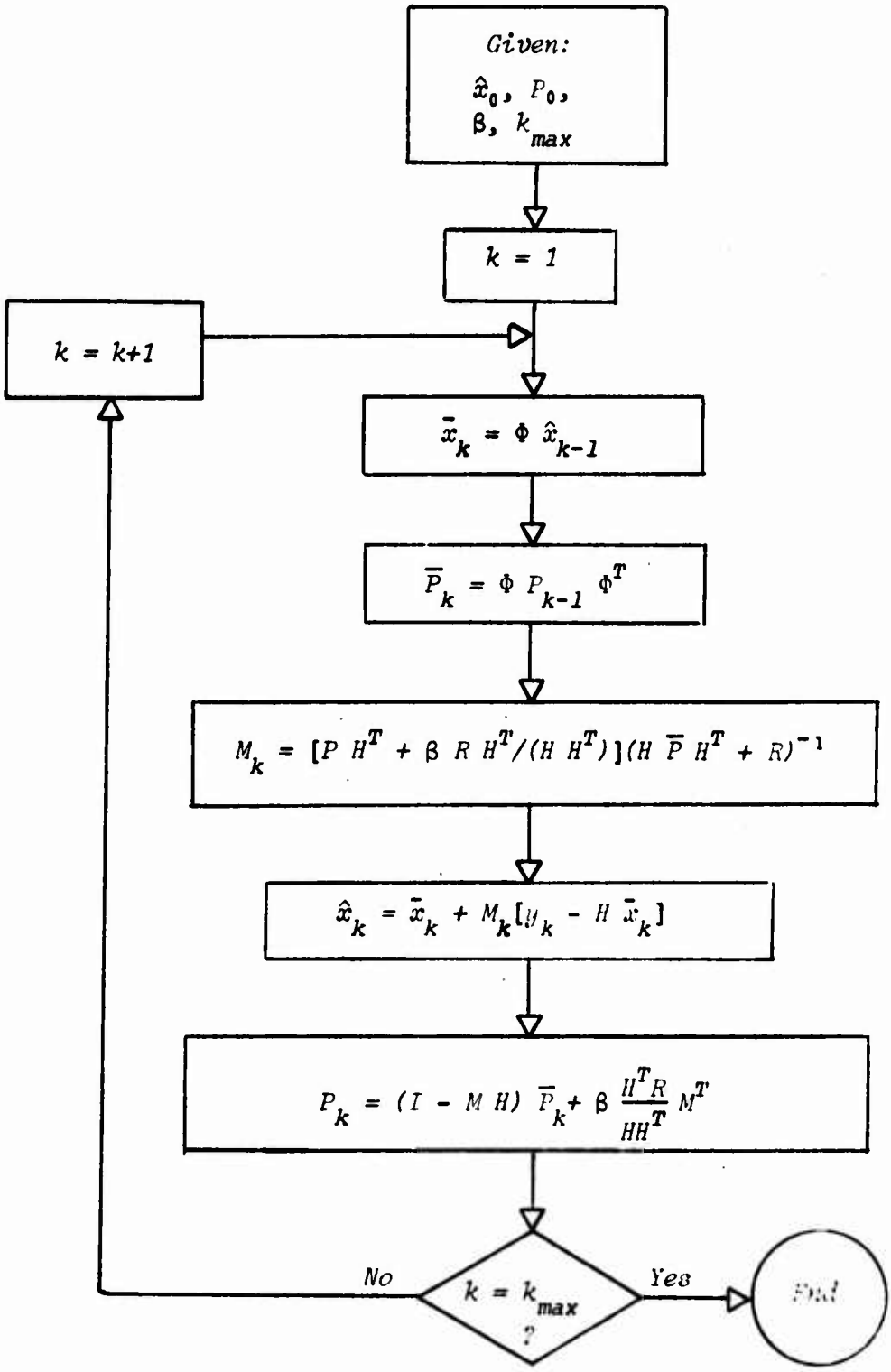


Figure 2.4 Schmidt Suboptimal Filter:
Additive Gain Term

Recalling that $x_k = \Phi_{k,m}\, x_m$,

$$J_k = \tfrac{1}{2}(x_k - \hat{x}_{k|m})^T\, \Phi_{k,m}^T\, P_{m|m}\, \Phi_{k,m} + \tfrac{1}{2} \sum_{i=m+1}^{k} (y_i - H_i\, \Phi_{i,k}\, x_k)^T\, R_i^{-1} (y_i - H_i\, \Phi_{i,k}\, x_k) \quad . \tag{2.4.2}$$

Further, define

$$\bar{y} = \begin{bmatrix} y_{m+1} \\ y_{m+2} \\ \cdot \\ \cdot \\ \cdot \\ y_k \end{bmatrix} \quad , \qquad \bar{H} = \begin{bmatrix} H_{m+1}\, \Phi_{m+1,k} \\ H_{m+2}\, \Phi_{m+2,k} \\ \cdot \\ \cdot \\ \cdot \\ H_k \end{bmatrix}$$

and

$$\bar{R}^{-1} = \begin{bmatrix} R_{m+1}^{-1} & & & \\ & R_{m+2}^{-1} & & \\ & & \ddots & \\ & & & R_k^{-1} \end{bmatrix} \quad . \tag{2.4.3}$$

Then,

$$J_k = \tfrac{1}{2}(x_k - \hat{x}_{k|m})^T\, P_{k|m}^{-1} (x_k - \hat{x}_{k|m}) + \tfrac{1}{2}(\bar{y} - \bar{H}\, x_k)^T\, \bar{R}^{-1} (\bar{y} - \bar{H}\, x_k) \tag{2.4.4}$$

Next take the gradient of J_k with respect to x_k, and set it equal to zero.

$$\delta J_k = P_{k|m}^{-1} (x_k - \hat{x}_{k|m}) - \bar{H}^T\, \bar{R}^{-1} (\bar{y} - \bar{H}\, x_k) = 0 \tag{2.4.5}$$

Solving for $x_k \equiv \hat{x}_{k|k}$ we obtain,

$$\hat{x}_{k|k} = [P_{k|m}^{-1} + \bar{H}^T\, \bar{R}^{-1}\, \bar{H}]^{-1}\, [\bar{H}^T\, \bar{R}^{-1}\, \bar{y} + \bar{P}_{k|m}^{-1}\, x_{k|m}]$$

or

$$\hat{x}_{k|k} = P_{k|k} [\bar{H}^T\, \bar{R}^{-1}\, \bar{y} + \bar{P}_{k|m}^{-1}\, x_{k|m}] \tag{2.4.6}$$

The fact that $P_{k|k} = [P_{k|m}^{-1} + \bar{H}^T \bar{R}^{-1} \bar{H}]^{-1}$ follows from equation (1.3.3).

Notice that if there is no *a priori* information about x_k then $P_{k|m}^{-1} = 0$ and *(2.4.6)* reduces to the standard weighted least squares algorithm. In this case, letting $N = k - m$, equation *(2.4.6)* is used to obtain *

$$\hat{x}_{k|(N)} = [\bar{H}^T \bar{R}^{-1} \bar{H}]^{-1} \bar{H}^T \bar{R}^{-1} \bar{y}$$

or

$$\hat{x}_{k|(N)} = P_{k|(N)} \bar{H}^T \bar{R}^{-1} \bar{y} \qquad (2.4.7)$$

Now combining equations *(2.4.6)* and *(2.4.7)* one obtains the following sequence:

$$\hat{x}_{k|k} = P_{k|k}[P_{k|(N)}^{-1} \hat{x}_{k|(N)} + P_{k|m}^{-1} \hat{x}_{k|m}]$$

$$P_{k|k}^{-1} \hat{x}_{k|k} = P_{k|(N)}^{-1} \hat{x}_{k|(N)} + P_{k|m}^{-1} \hat{x}_{k|m}$$

$$\hat{x}_{k|(N)} = P_{k|(N)}[P_{k|k}^{-1} \hat{x}_{k|k} - P_{k|m}^{-1} \hat{x}_{k|m}] \qquad (2.4.8)$$

with

$$P_{k|(N)} = [P_{k|k}^{-1} - P_{k|m}^{-1}]^{-1} \qquad (2.4.9)$$

The limited memory filter equations described above are obviously more complex to implement than the previously presented filtering algorithms. Essentially two Kalman sequences are required to be run for each "batch" of N observations, and three matrix inverses are required every N observations. Further, the limited memory filter estimate is obtained only every N

---

* $\hat{x}_{k|(N)}$ is the state estimate at t_k based on observations to t_k, less those to t_m, i.e., on the batch of the last $N = k - m$ observations. The same meaning applies to the error covariance, $P_{k|(N)}$.

observations. This number may be increased, but only at the expense of a larger number of inverses of $P_{k|k}$, $P_{k|m}$, and $P_{k|(N)}$.

It is possible to modify Jazwinski's algorithm in order to eliminate one of the matrix inversions. By factoring out $P_{k|k}$, equation *(2.4.9)* may be written as

$$P_{k|(N)} = [I - P_{k|k} P_{k|m}^{-1}]^{-1} P_{k|k} \qquad (2.4.10)$$

Also, solving *(2.4.9)* for $P_{k|k}^{-1}$ and substituting into *(2.4.8)* yields

$$\hat{x}_{k|(N)} = \hat{x}_{k|k} + P_{k|(N)} P_{k|m}^{-1}[\hat{x}_{k|k} - \hat{x}_{k|m}] \qquad (2.4.11)$$

Thus the inversion of $P_{k|k}$ has been eliminated, and *(2.4.10)* and *(2.4.11)* are the equations to be employed. It is, of course, necessary to obtain $\hat{x}_{k|k}$ and hence $P_{k|k}$ is required.

The limited memory filter is implemented as a sequential estimator. Figure 2.5 diagrams the modified procedure which uses *(2.4.10)* and *(2.4.11)*. The Kalman filter equations are run from $k = 0$ to $k = m = N$, processing observations and obtaining estimates in the usual manner. This represents an initialization phase. The values of $P_{m|m}$ and $\hat{x}_{m|m}$ are stored. The Kalman equations are then run from $k = m$ to $k = m + N$, also in the usual manner, and the state transition matrix $\Phi_{m+N,m}$ is obtained using $\Phi_{k+2,k} = \Phi_{k+2,k+1}\ \Phi_{k+1,k}$. The predicted values, $P_{m+N|m}$ and $\hat{x}_{m+N|m}$ are found using $\Phi_{m+N,m}$ and equations *(2.4.10)* and *(2.4.11)* applied to obtain $\hat{x}_{m+N|(N)}$ and $P_{m+N|(N)}$. The process is then repeated (except for the initialization phase) using the limited memory filter outputs as the inputs for the next cycle (from $m = 2N$ to $m = 3N$).

It appears from examination of the equations for finding $P_{k|(N)}$, *(2.4.9)* or *(2.4.11)*, that numerical problems could be encountered, particularly where short word length computers must be used. Equations *(2.4.9)* and

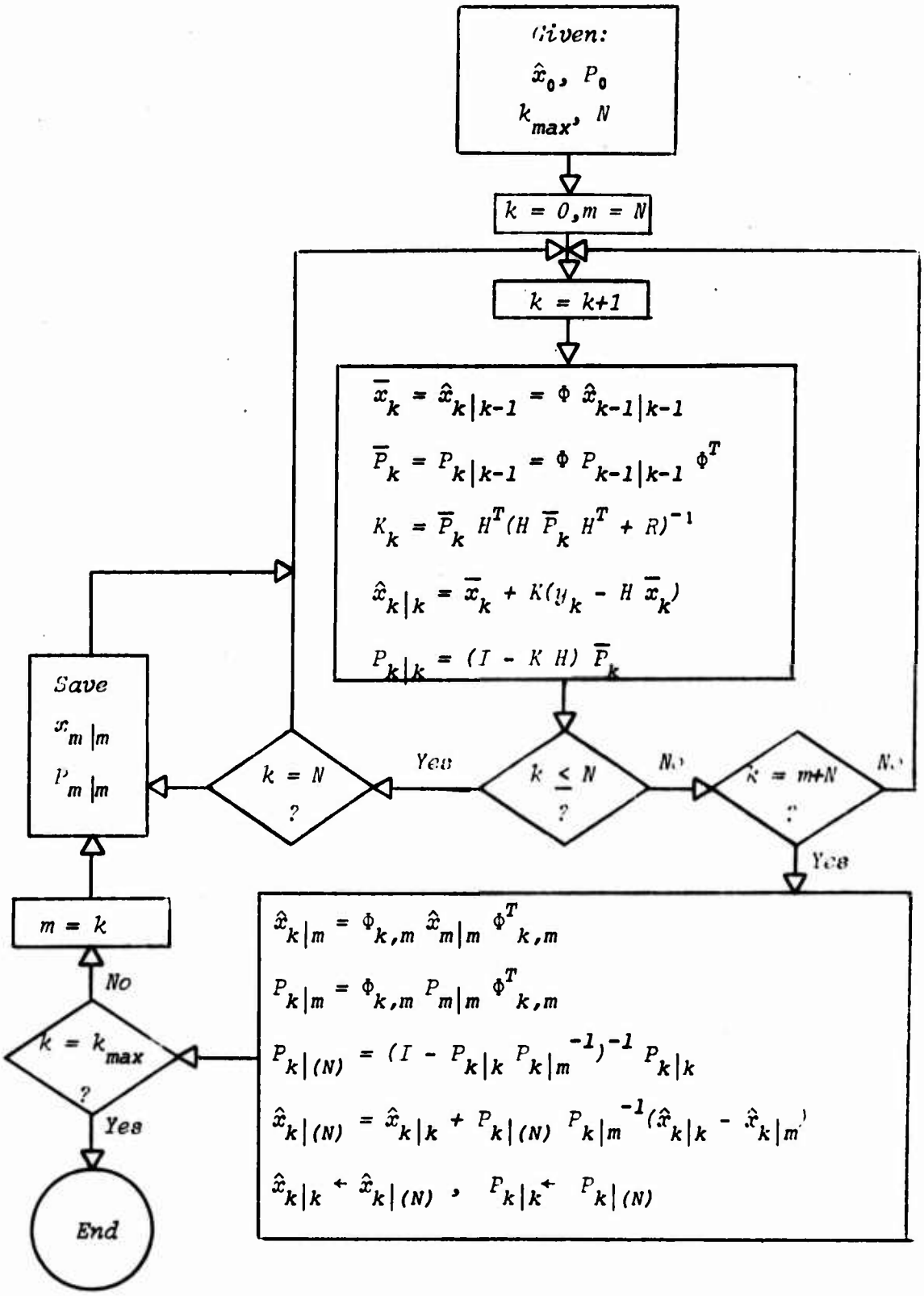


Figure 2.5 Modified Limited Memory Filter

(2.4.10) require computation of the inverse of the difference of the inverses of two matrices of similar values.

It is possible to offset this state of affairs. An alternative due to Jazwinski (5, 258), reduces the number of P-matrix inversions to one by working directly with the inverses only. Equation *(1.3.9)* from section 1.3 provides the necessary recursion relation. Dropping the m-subscripts from *(1.3.9)* one obtains at time t_k, based on measurements to t_{k-1},

$$P^{-1}_{k|k-1} = \Phi^{-T}_{k,k-1} P^{-1}_{k-1|k-1} \Phi^{-1}_{k,k-1} + H_k^T R_k^{-1} H_k \qquad (2.4.12)$$

For prediction only, the information term, $H^T R^{-1} H$, for time t_k is ignored. Thus at t_k, based on measurements to t_m,

$$\bar{P}^{-1}_{k|m} = \Phi^{-T}_{k,m} P^{-1}_{m|m} \Phi^{-1}_{k,m} \qquad (2.4.13)$$

This, of course, requires taking the inverse of Φ. If Φ is symplectic*, its inverse is readily obtained. Alternatively, instead of integrating $\dot{\Phi} = A\,\Phi$, the inverse may be found by direct integration. $\dot{\Phi}^{-1}$ is found as follows:

$$\Phi^{-1}\,\Phi = I$$

$$\dot{\Phi}^{-1}\,\Phi + \Phi^{-1}\,\dot{\Phi} = 0$$

$$\dot{\Phi}^{-1} = -\Phi^{-1} A$$

It should be apparent that all the previously presented algorithms may be implemented such that $\bar{x}$ and $\bar{P}$ are obtained by integrating their

---

* An even-dimensioned matrix, Φ, is said to be symplectic if

$$\Phi^T J\,\Phi = J, \text{ where } J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

Post-multiplying by Φ^{-1} and pre-multiplying by J gives

$$\Phi^{-1} = -J\,\Phi^T J$$

appropriate differential equations. Similarly, differential equations may be obtained for the prediction and update of P^{-1}. First,

$$P^{-1} P = I$$

$$\dot{P}^{-1} P + P^{-1} \dot{P} = 0$$

$$\dot{P}^{-1} = -P^{-1} \dot{P} P^{-1}$$

and using equation *(1.2.39)* with $Q = 0$, one obtains for the prediction,

$$\dot{P}^{-1}(t) = -P^{-1}(t) A(t) - A^T(t) P^{-1}(t) \qquad (2.4.14)$$

The differential equation for updating P^{-1} may be found by applying a limiting process to *(2.4.12)* similar to that used in obtaining equation *(1.2.39)*. First,

$$P^{-1}(\tau) = \Phi^{-T}(\tau, t) P^{-1}(t) \Phi^{-1}(\tau, t) + H^T(\tau) R^{-1}(\tau) H(\tau)$$

$$= [\Phi^{-T}(t, t) + \dot{\Phi}^{-T}(t, t) \Delta t + \ldots] P^{-1}(t)$$

$$\cdot [\Phi^{-1}(t, t) + \dot{\Phi}^{-T}(t, t) \Delta t + \ldots] + H^T(\tau) R^{-1}(\tau) H(\tau)$$

Carrying out the indicated multiplication,

$$P^{-1}(\tau) = P^{-1}(t) - A^T(t) P^{-1}(t) \Delta t - P^{-1}(t) A(t) + \ldots + H^T(\tau) R^{-1}(\tau) H(\tau)$$

Performing the limiting process requires subtracting $P^{-1}(t)$ from both sides, and dividing by Δt. However, when taking the limit as $\Delta t \to 0$ the information term becomes infinite implying that the observations are perfect (no observation noise). This is inconsistent with the postulated observation process, hence we replace* $R(\tau)$ with $R(\tau)/\Delta t$. With this change, the limiting process is applied to obtain

$$\dot{P}^{-1}(t) = -A^T(t) P^{-1}(t) - P^{-1} A(t) + H^T(t) R^{-1}(t) H(t) \qquad (2.4.15)$$

---

* Further rationale and a discussion of a white noise process as the limit of a white noise sequence is given in (5, 83-84).

with

$$P^{-1}(0) = P_0^{-1}$$

When the information term, $H^T R^{-1} H$, is added at discrete times, the approach used is to predict to $t = t_k$ using *(2.4.14)* and then add the information terms at t_k, i.e.,

$$P_{k|k-1}^{-1} = -\int_{t_{k-1}}^{t_k} [P^{-1}(t)\, A(t) + A^T(t)\, P^{-1}(t)]\, dt$$

$$P_{k|k}^{-1} = P_{k|k-1}^{-1} + H_k^T R_k^{-1} H_k \qquad (2.4.16)$$

Reverting to the notation associated with the filter, there results

$$P_{k|m}^{-1} = -\int_{t_m}^{t_k} [P^{-1}(t|t_m)\, A(t) + A^T(t)\, P^{-1}(t|t_m)]\, dt \qquad (2.4.17)$$

$$P_{k|(N)}^{-1} = \sum_{j=m}^{k-1} \left\{ -\int_{t_j}^{t_{j+1}} [P^{-1}(t|t_{j-m})\, A(t) + A^T(t)\, P^{-1}(t|t_{j-m})]\, dt \right.$$

$$\left. + H^T(j+1)\, R^{-1}(j+1)\, H(j+1) \right\}$$

with

$$P^{-1}(t_m|0) = 0 \qquad (2.4.18)$$

As before, it is not necessary to compute $P_{k|k}^{-1}$ (although $P_{k|k}$ is required) since equation *(2.4.11)* provides the limited memory estimate. Thus only one P-matrix inversion is required: that of $P_{k|(N)}^{-1}$ to obtain $P_{k|(N)}$. However, the inverse of the state transition matrix is required, or else *(2.4.17)* and *(2.4.18)* must be used. If Φ^{-1} cannot be easily obtained, the advantage of this approach is somewhat decreased. The implementation of the alternate limited memory filter is diagramed in Figure 2.6.

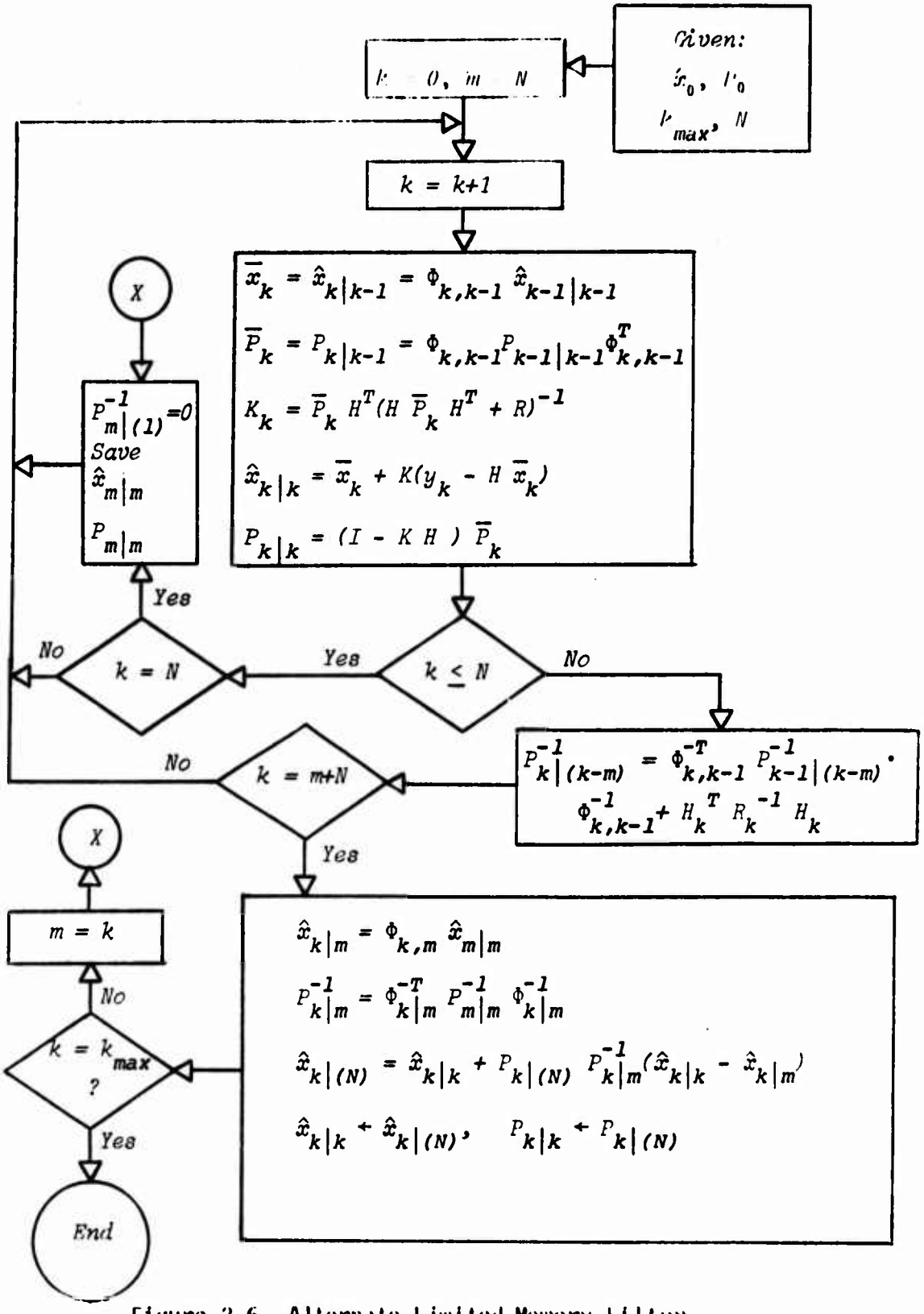


Figure 2.6 Alternate Limited Memory Filter

2.5 Summary

In this chapter the non-adaptive error compensation algorithms of age-weighting, gain-scaling, additive gain term, and limited memory filtering have been presented. Some of the original algorithms have been extended to provide alternate forms, and relationships derived showing certain equivalences among the first three.

Equations *(2.1.1)* give the basic age-weighting algorithm as conceived by Tarn and Zaborsky, with Fagin's form given by *(2.1.1)*.

Schmidt's gain-scaling algorithm is given by equations *(2.2.1)* and *(2.2.3)*. The relationship between the scaling factor, b, and ε $(=HPH^T)$ is given by *(2.2.5)*. Modifications of the algorithm are given by the alternate equations for determining b, equations *(2.2.11)* and *(2.2.14)*.

The second filter design of Schmidt's adds a judiciously selected term to the optimal gain, K. The resulting gain and covariance are given by equations *(2.3.2)* and *(2.3.5)* or *(2.3.6)*. The weighting factor, β, associated with the additional gain term is shown to be related to ε by equation *(2.3.9)*, and related to b by *(2.3.7)*.

Jazwinski's limited memory filter is given by equations *(2.4.8)* and *(2.4.9)* together with the Kalman filtering algorithm. A modification to the basic limited memory filter eliminates one of the three matrix inverses, and is given by *(2.4.10)* and *(2.4.11)*. An alternate form employs the inverse of the error covariance directly, thereby reducing the number of matrix inverses to one, and is particularly advantageous if the state transition matrix is symplectic.

In the next chapter we turn our attention to adaptive techniques, both optimal and suboptimal.

Chapter 3

ADAPTIVE METHODS

3.1 Estimating the State Noise Covariance

When modeling errors are assumed to be represented as purely random, uncorrelated noise, it is often a problem to determine the appropriate value of Q, the state noise covariance. As indicated in section 1.5, there is a preponderance of adaptive techniques which may be used to estimate the appropriate noise covariances. Most of these, particularly Bayesian and maximum likelihood approaches, have been thoroughly investigated (25), (26), (27), (28). In this section an adaptive method due to Jazwinski is presented which offers a reasonably simple and easily implemented scheme for adaptively estimating the state noise covariance matrix. While the algorithm is derived using a maximum likelihood approach, an interesting property is that in a simplified form it becomes just the covariance matching technique.

Consider the following predicted residuals, called the innovations (23)

$$\bar{r}_k = y - H_k \bar{x}_k = H_k \tilde{x}_k + v_k, \quad E\{v_k\} = 0, \quad E\{v_k v_j^T\} = R_k \delta_{kj} \qquad (3.1.1)$$

For simplicity, $\bar{r}_k$ is assumed to be restricted to the scalar case. Forming the covariance of $\bar{r}_k$ one obtains

$$E\{\bar{r}_k^2\} = H_k \bar{P}_k H_k^T + R_k \qquad (3.1.2)$$

since $E\{\tilde{x}_k v_k^T\} = 0$. $\bar{P}$ is given by equation (1.2.17). It can be shown (23) that the innovations given by (3.1.1) are uncorrelated, i.e.,

$$E\{\bar{r}_k \bar{r}_j\} = 0, \quad j > k,$$

and that $\bar{r}_k$ is a Gaussian white noise sequence. Thus the joint probability density function of the innovations sequence $\bar{r}_{k+1}, \bar{r}_{k+2}, \ldots, \bar{r}_{k+N}$ is

$$p(\bar{r}_{k+1}, \bar{r}_{k+2}, \ldots, \bar{r}_{k+N}) = p(\bar{r}_{k+1}) \cdot p(\bar{r}_{k+2}) \cdot \ldots p(\bar{r}_{k+N}) \qquad (3.1.3)$$

where

$$p(\bar{r}_{k+j}) = \frac{1}{\sqrt{2\pi}} \left[H_{k+j} \bar{P}_{k+j} H_{k+j}^T + R_{k+j} \right]^{-\frac{1}{2}}$$

$$exp \left\{ -\tfrac{1}{2} \left[\frac{\bar{r}_{k+j}}{H_{k+j} \bar{P}_{k+j} H_{k+j} + R_{k+j}} \right] \right\} \qquad (3.1.4)$$

Suppose during operation of the filter equations (say, Algorithm I) that $t = t_k$, and the predicted values $\bar{P}_{k+1}$ and $\bar{x}_{k+1}$ are to be computed. For N predicted residuals the object is to find $Q_{k|N}$ (the state noise covariance at t_k based on N innovations) such that (3.1.3) is maximized. In other words, $Q_{k|N}$ is that value of Q which yields the most likely innovations sequence $\bar{r}_{k+1}, \bar{r}_{k+2}, \ldots \bar{r}_{k+N}$. Thus very large values of Q imply less likely sequences of the innovations.

Consider the maximization of (3.1.3) for the case of one residual. First, the likelihood function is chosen as the joint density (3.1.4), and its logarithm is taken to facilitate the maximization (since p and its logarithm are monotonic, extremizing $ln\ p$ extremizes p).

$$J = ln\ p(\bar{r}_{k+1}) = ln(\tfrac{1}{2\pi})^{\frac{1}{2}} - ln\ (H_{k+1} \bar{P}_{k+1} H_{k+1}^T + R_{k+1})$$

$$- \tfrac{1}{2} \left(\frac{\bar{r}_{k+1}^2}{H_{k+1} \bar{P}_{k+1} H_{k+1}^T + R_{k+1}} \right) \qquad (3.1.5)$$

Taking the variation with respect to Q gives

$$\delta J = \left\{ -\tfrac{1}{2} \frac{1}{(H \bar{P} H^T + R)} + \tfrac{1}{2} \frac{\bar{r}^2}{[H \bar{P} H^T + R]^2} \right\} H \delta \bar{P} H^T \qquad (3.1.6)$$

and assuming $H\delta\bar{P}H^T > 0$, the necessary condition $\delta J = 0$ yields the equation

$$\bar{r}^2 - H \bar{P} H^T - R = 0 \qquad (3.1.7)$$

The sufficient condition for a maximum is $\delta^2 J < 0$; hence,

$$\left\{ \tfrac{1}{2} \frac{1}{(H \bar{P} H^T + R)^2} - \frac{\bar{r}^2}{[H \bar{P} H^T + R]^3} \right\} < 0$$

or

$$\bar{r}^2 > \tfrac{1}{2} [H \bar{P} H^T + R] \qquad (3.1.8)$$

which is met if *(3.1.7)* holds. Next, using *(1.2.17)*,

$$H_{k+1} \Gamma_{k+1,k} Q_{k|1} \Gamma^T_{k+1,k} H^T_{k+1} =$$

$$\bar{r}^2_{k+1} - H_{k+1} \Phi_{k+1,k} P_k \Phi^T_{k+1,k} H^T_{k+1} - R_{k+1} \qquad (3.1.9)$$

This equation is simply the equation for the covariance matching approach to estimating Q based on one residual, $\bar{r}_{k+1}$. In evaluating *(3.1.9)* the actual residual given by equation *(3.1.1)* is used. Further, for Q to be non-negative definite, the right hand side must be positive. Otherwise $Q_{k|1}$ is set to zero. Unfortunately, $H\Gamma$ is usually not invertable, hence, resort must be made to a pseudo-inverse or some other technique. Defining, for convenience,

$$\Lambda \equiv H \Gamma$$

the pseudo inverse of Λ (45, 82-89)

$$\Lambda^{\#} = \frac{\Lambda^T}{\Lambda \Lambda^T}$$

yields

$$Q_{k|1} = \frac{\Lambda^T}{\Lambda\Lambda^T} [\bar{r}^2_{k+1} - H \Phi P_k \Phi^T H^T - R] \frac{\Lambda}{\Lambda\Lambda^T} \qquad (3.1.10)$$

Equation *(3.1.10)* does not provide a unique solution for Q. However, in the special case where Q is a scalar, say q, one does obtain a unique solution:

$$q_{k|1} = \frac{\bar{r}^2 - H \Phi P \Phi^T H^T - R}{\Lambda \Lambda^T} \qquad (3.1.11)$$

Here again we require the right hand side to be non-negative. A disadvantage is that one residual is hardly a sound statistical sample upon which to base an estimate of Q. A *post hoc* modification found to be effective is to employ the sample mean of the square of N predicted residuals,

$$\gamma_N = \frac{1}{N} \sum_{j=0}^{N-1} \bar{r}_{m+j}^{\,2}, \quad m = k + 1 \qquad (3.1.12)$$

in place of $\bar{r}_{k+1}^{\,2}$. This requires storing N observations, which would not be so bad in itself, but the state estimate then either periodically lags the observations by N points, or else the last N estimates must be reprocessed. Either way the result is the same, but the former is computationally more efficient since the state estimate is obtained only once for each time point.

Another modification which is computationally easier to employ and which does not severly degrade the accuracy of the algorithm is to use the sample mean of the history of the last N predicted residuals squared. In this case,

$$\gamma_N = \frac{1}{N} \sum_{j=0}^{N-1} \bar{r}_{m-j}^{\,2} \qquad (3.1.13)$$

In this approach the filter equations are operated for the first m observations, storing the predicted residual at each time point, t_{k+1}. For each $m \leq N$, the sample mean,

$$\gamma_m = \frac{1}{m} \sum_{j=0}^{m-1} \bar{r}_{m-j}^{\,2}, \quad m = 1, 2, \ldots, N \qquad (3.1.14)$$

is formed and used in the estimation of Q. For each $m > N$ a new squared residual is added, and the oldest one is discarded. In this way a "moving window" viewing the most recent N squared residuals is maintained. This modified approach is diagrammed in Figure 3.1.

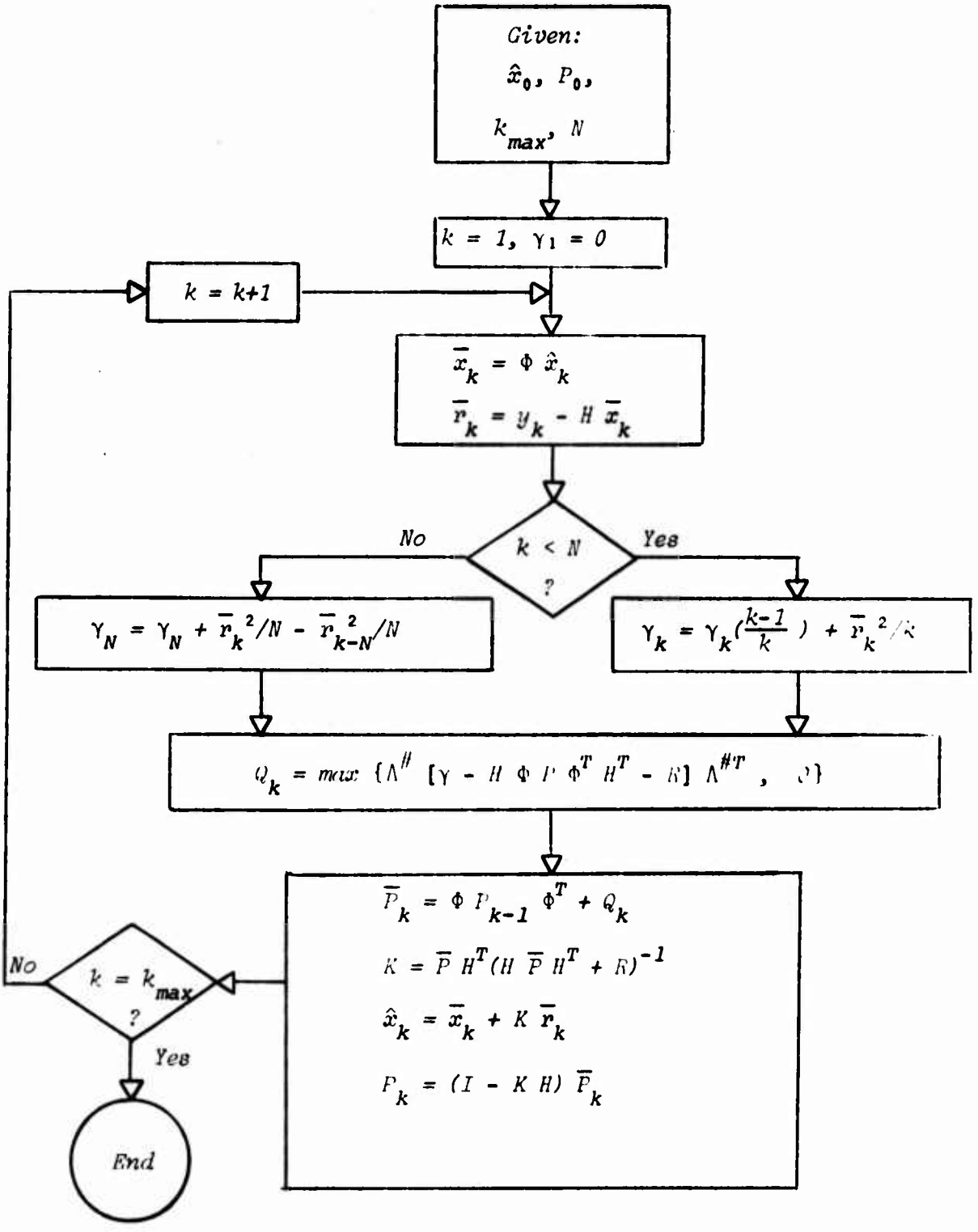


Figure 3.1 Adaptive Estimation of Q

Note that both of the approaches reduce to the optimal case of one predicted residual when $N = 1$. In this case *(3.1.12)* and *(3.1.13)* are identical.

Further modifications are possible to improve the single residual estimator (5, 314); however, these tend to become rather complex. Also, when more than one residual is used in *(3.1.3)*, the resulting set of equations to be solved is nonlinear. In fact, they are basically the types of equations found in other adaptive filtering methods (26), (28).

3.2 Adaptive Estimation of Suboptimal Filter Parameters

Since the suboptimal methods of Chapter 2 effectively alter the state error covariance, it is reasonable to assume that the adaptive algorithm of section 3.1 can be applied to estimate the various suboptimal filter parameters. In each case, relating the parameters to the state noise covariance provides the necessary link to allow the adaptive estimation of the parameters.

For the data age-weighting technique, we can equate the standard error covariance prediction equation *(1.2.17)* to the age-weighted error covariance prediction equation (with $Q = 0$):

$$\bar{P}_{k+1} = \Phi P_k \Phi^T + \Gamma Q_k \Gamma^T = \Phi s P \Phi^T \qquad (3.2.1)$$

Using this relation in *(3.1.7)* gives

$$\bar{r}_{k+1}^2 - H_{k+1} \Phi s P_k \Phi^T H_{k+1} - R_{k+1} = 0$$

or solving for the scalar, s,

$$s = \frac{\bar{r}^2 - R}{H \Phi P \Phi^T H} \qquad (3.2.2)$$

Again the *post hoc* modification is made and γ_N given by *(3.1.19)* or *(3.1.18)* is utilized.

$$a = \frac{\gamma_N - R}{H \Phi P \Phi^T H^T} \qquad (3.2.3)$$

Since a is required to be greater than or equal to 1, (3.2.3) may not be employed until

$$\gamma_N \geq H \Phi P \Phi^T H^T + R \qquad (3.2.4)$$

To adaptively estimate the gain scaling parameter, b, in section 2.2, the optimal and suboptimal gain expressions are equated.

$$b P' H^T (H P' H^T + R)^{-1} = (P' + Q') H^T (H P' H^T + H Q' H^T + R)^{-1} \qquad (3.2.5)$$

where $P' \equiv \Phi P \Phi^T$ and $Q' \equiv \Gamma Q \Gamma^T$. Solving for $H Q' H^T$ one obtains, after premultiplying by H,

$$b H P' H^T (H P' H^T + H Q' H^T + R) = H(P' + Q') H^T (H P H^T + R)$$

$$H Q' H^T = H \Gamma Q_k \Gamma^T H^T = \frac{(b-1) H P' H^T (H P' H^T + R)}{[R - (b-1) H P' H^T]} \qquad (3.2.6)$$

Substituting into equation (3.1.9),

$$\frac{(b-1) H P' H^T (H P' H^T + R)}{R - (b-1) H P' H^T} = \gamma_N - H P' H^T - R$$

which, solved for b, yields

$$b = \frac{(\gamma_N - R) [H \Phi P \Phi^T H^T + R]}{H \Phi P \Phi^T H^T} \geq 1 \qquad (3.2.7)$$

where the inequality on the right is a condition on b established in section 2.2. Also it is noted that equation (3.2.7) requires $\gamma_N > R$, a condition on its use.

The modification (2.2.11) to the gain scaling technique may be formulated similarly to adaptively estimate its associated parameter, α. Equating (2.2.12) with the gain equation gives

$$\frac{P^- H^T}{H P^- H^T}\,\alpha = (P^- + Q^-)\,H^T(H P^- H^T + H Q^- H^T + R)^{-1} \qquad (3.2.8)$$

which, when premultiplied by H and solved for $H Q^- H^T$ yields

$$H Q^- H^T = \frac{\alpha R - (1 - \alpha) H P^- H^T}{(1 - \alpha)} \qquad (3.2.9)$$

Substituting into *(3.1.9)* and solving for α yields the simple equation,

$$\alpha = 1 - R/\gamma_N \qquad (3.2.10)$$

Conditions on α require that *(3.2.10)* produce an $\alpha \geq H K$.

For the modification given by *(2.2.14)* the appropriate adaptive form is similarly found. Equating the gains,

$$(P^- + Q^-)\,H^T(H P^- H^T + H Q^- H^T + R)^{-1} =$$

$$(P^- H^T + \beta R P H^T/H P^- H^T)\,(H P^- H^T + R)^{-1} \qquad (3.2.11)$$

premultiplying by H and solving for $H Q^- H^T$ gives

$$H P^- H^T(H P^- H^T + R) + H Q^- H^T(H P^- H^T + R) =$$

$$H P^- H^T(H P^- H^T + R) + H P^- H^T H Q^- H^T + \beta R(H P^- H^T + R) + \beta^2 H Q^- H^T$$

$$H Q^- H^T R(1 - \beta) = \beta R(H P^- H^T + R)$$

$$H Q^- H^T = H \Gamma Q \Gamma^T H^T = \frac{\beta(H P^- H^T + R)}{(1 - \beta)}$$

(As before, $P^- = \Phi P \Phi^T$ and $Q^- = \Gamma Q \Gamma^T$.) Substituting into *(3.1.9)* and solving for β one obtains

$$\beta = 1 - \frac{(H P^- H^T + R)}{\gamma_N} = 1 - \frac{[H \Phi P \Phi^T H^T + R]}{\gamma_N} \qquad (3.2.12)$$

with the conditions on β that $0 \leq \beta \leq 1$ and hence

$$\gamma_N \geq H \Phi P \Phi^T H^T + R$$

Finally we note that for the additive gain term technique, a similar procedure to determine the adaptive form yields results identical with equation *(3.2.12)*. This is true because equating gains and premultiplying by H gives $H\ \beta\ R\ H^T/H\ H^T = \beta\ R$, and from then on the equations are identical.

The previously discussed suboptimal adaptive algorithms are easily implemented in accordance with Figures 3.2, 3.3, and 3.4. Either of the predicted residual sample means discussed in section 3.1 may be used; hence their computation is not shown in these figures.

3.3 Estimation of Auto-Correlated Model Errors

One of the most effective and useful ways to compensate for dynamic modeling errors is to estimate such biases directly, including them as part of the state vector. In general, model errors are not purely random, but have time correlated components as well. Thus, by assuming some functional form for these components, their values may be estimated from observation point to observation point.

There are several advantages in this approach. First, of course, is that the estimation accuracy is improved since model errors are compensated. Second, insight into the nature of the modeling errors is obtained as an additional benefit, and through off-line data analysis, the form of the dynamic model can be refined for future use. Third, since the state vector is simply augmented, the estimation equations remain essentially unchanged from the classical Kalman sequential form. Of course, lengthening the state vector increases the computational load; however, this is not often a detriment considering the state of present computer technology.

A further motivation for the adaptive estimation of model errors should be noted. It is true that small bounded errors may be effectively

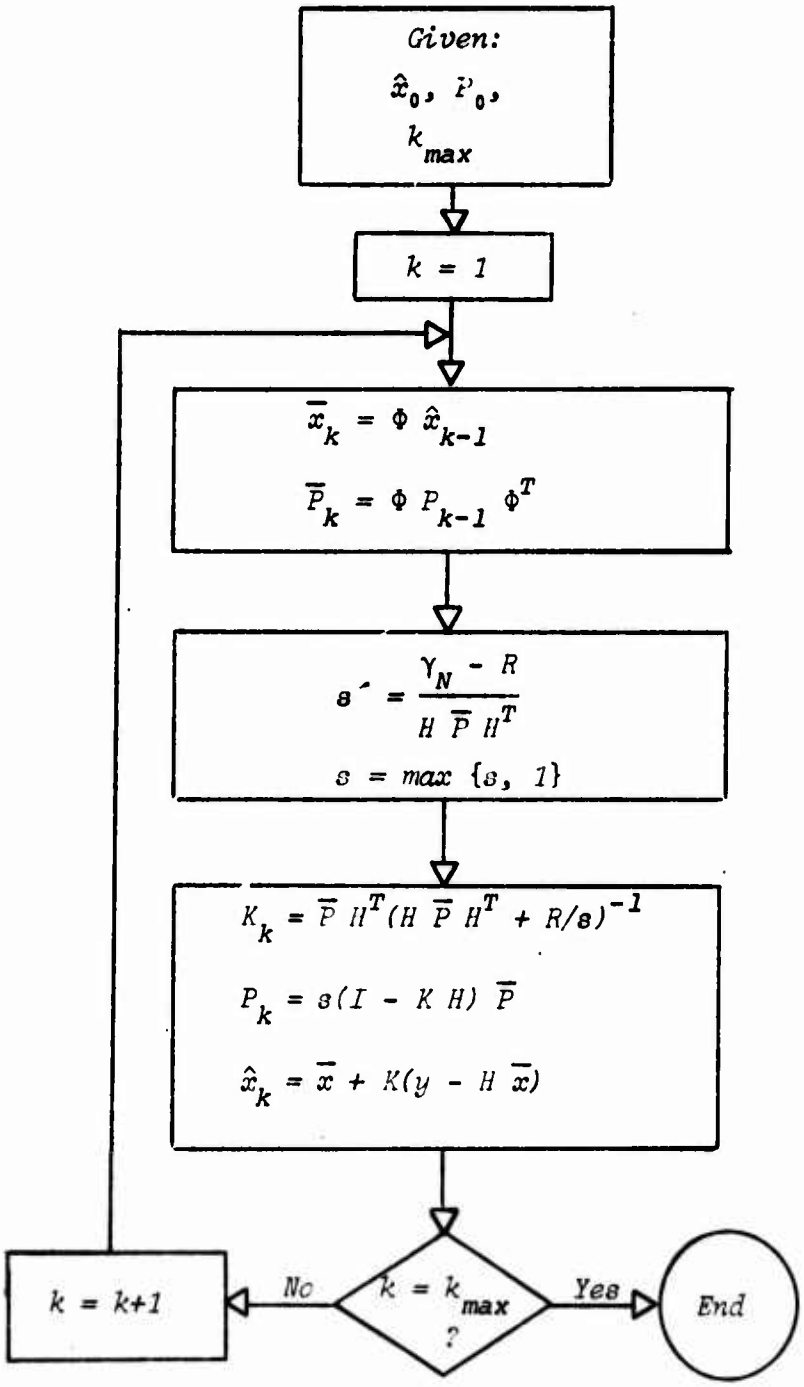


Figure 3.2 Adaptive Age-Weighting Filter

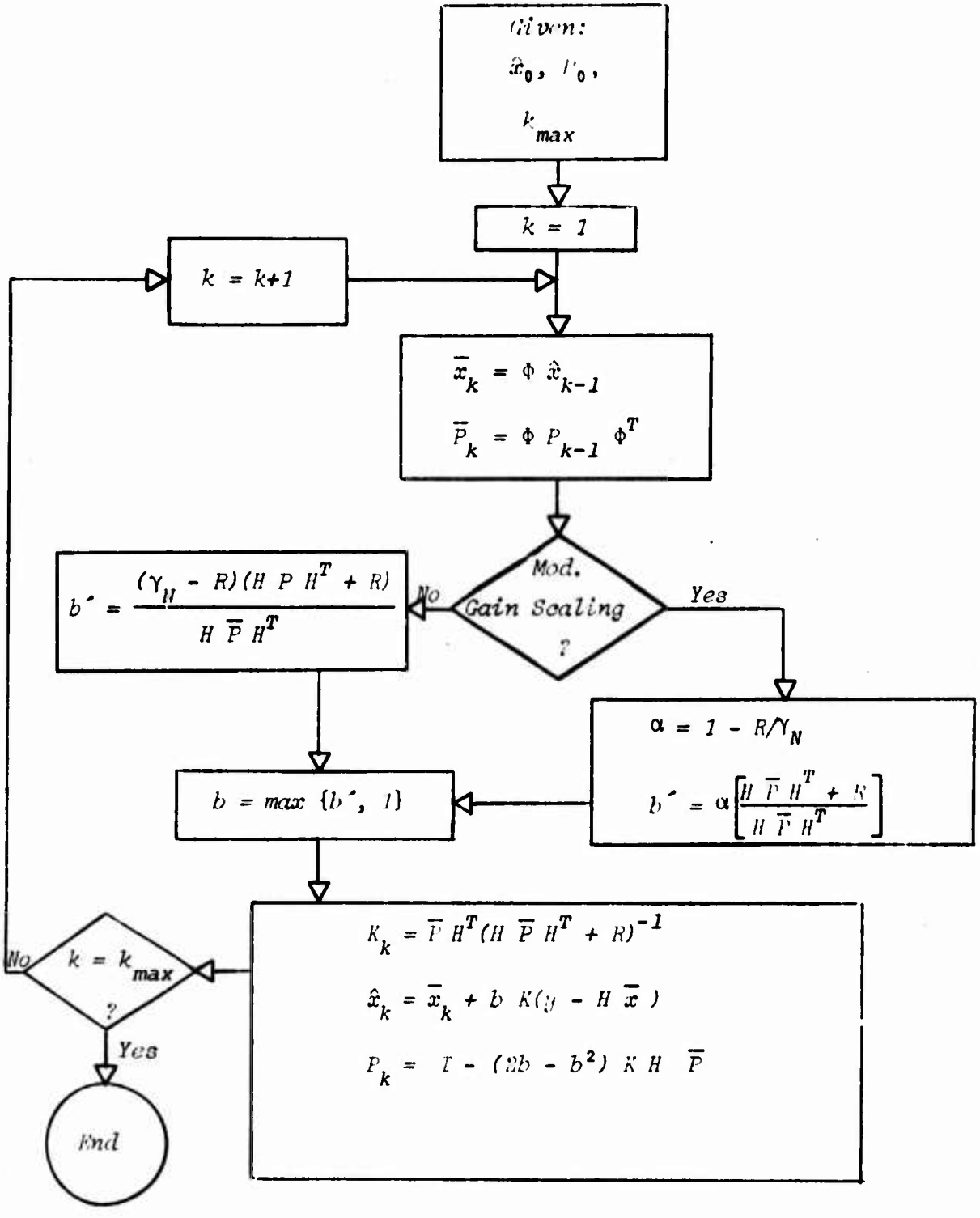


Figure 3.3 Adaptive Scaled Gain Filter

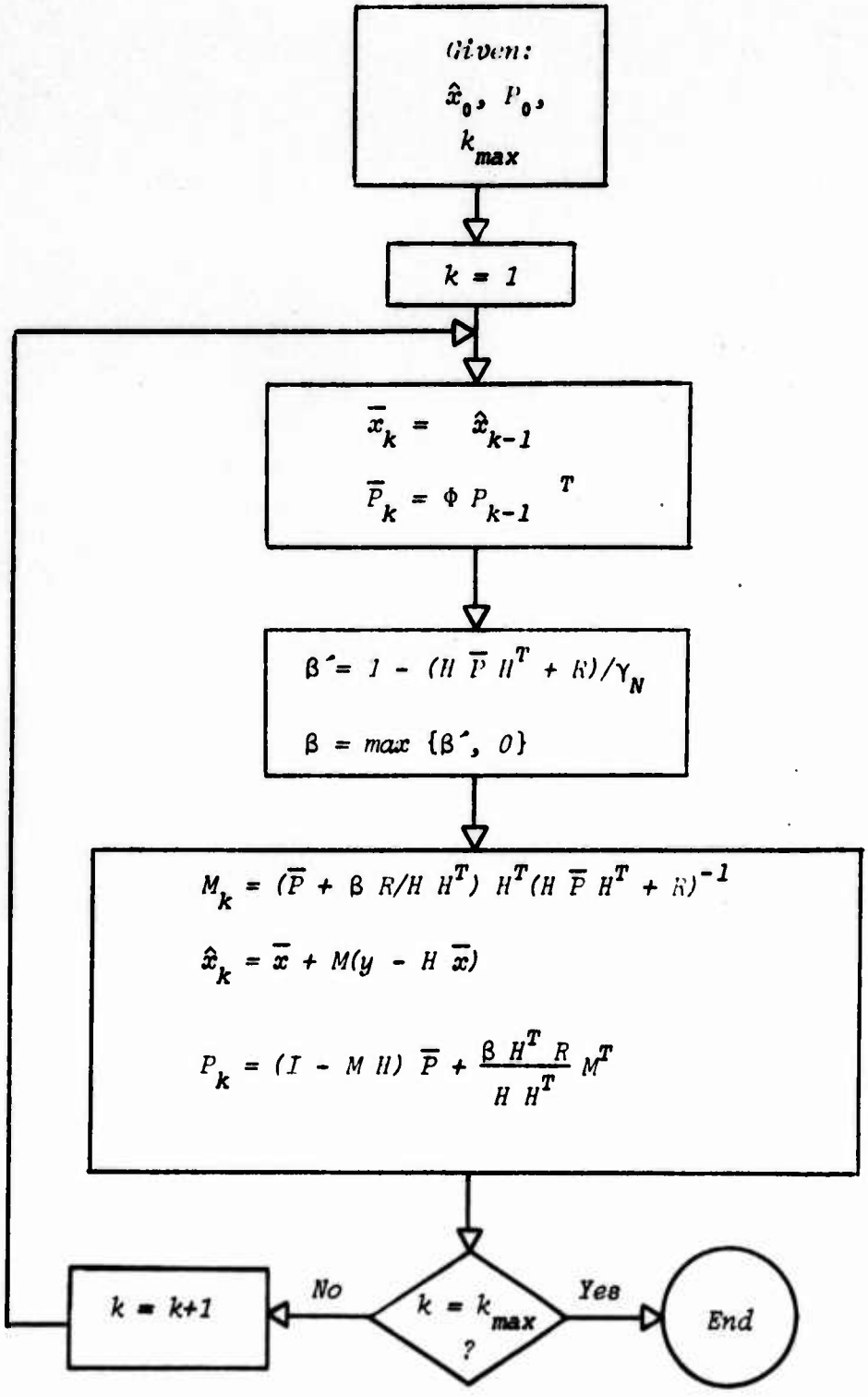


Figure 3.4 Adaptive Additive Gain Filter

compensated (as long as filter operation is within the range of linearity of the system dynamics) by using a "state noise" approach, i.e., simply determining the Q-matrix. However, for large or unbounded errors such an approach is usually impractical if not impossible. Thus extracting as much of the auto-correlated error as possible from the model uncertainties results in a smaller, purely random component. Consequently, the associated state noise covariance matrix, Q, is less, and since this establishes a minimum upon the state error covariance matrix, a smaller Q offers greater confidence in the state estimate.

The question concerning the type of functions to be used in modeling the correlated errors is really an open one. The answer depends upon the application where two important considerations are the filter operating time and the sampling interval (more will be said of this later). Approaches have included the use of simple power series in time, Tchebycheff polynomials, and Fourier series. For our purposes, the term structure is introduced to mean the functional form of the correlated model errors.* Thus the structure may be linear or nonlinear according to the form of the model error approximating functions. Incidentally, the case of purely random errors is denoted here as a null structure -- no time correlated components. Usually, the assumed error form is some linear combination of a linear or nonlinear structure and a null structure, i.e., of a time correlated component and a purely random component.

The incorporation of model error structures into the estimation process can occur in basically one of two explicit ways. Denoting the vector of

---

* Lainiotis (38) uses structure to mean the dimensionality of the state vector. Our definition may be seen to include this aspect.

error components by e, one can write a differential form,

$$\dot{e}(t) = g(e, a, t) + h(t)\, u_e(t)$$

$$\dot{a}(t) = 0 \qquad (3.3.1)$$

or an algebraic form, $e(t) = G(e_0, a, t) + H(t)\, w_e(t)$ with the accompanying differential form

$$\dot{e}_0(t) = 0$$

$$\dot{a}(t) = 0 \qquad (3.3.2)$$

where a is a parameter vector of constants, and e_0 is the initial condition vector, $e(t_0)$, and w_e is state noise with covariance $Q(t)$. $u_e(t)$ is zero mean, uncorrelated noise with covariance* $E\{u_e(t)\, u_e^T(s)\} = Q_e(t)\, \delta(t - s)$. Which form is used is of no particular importance in the subsequent discussion. If a closed form solution, or even an approximate solution, is available for the expected value of (3.3.2), an easier implementation may be possible since the corresponding components of the state transition matrix are simply the identity matrix of appropriate dimension. If no explicit solution of $E\{\dot{e}\}$ is available, clearly (3.3.1) is to be used. Further, if $g(e, a, t)$ is nonlinear in e and a, an appropriate linearization must be made in order to use the estimation equations described in section 1.2.

Since the model errors are estimated as part of the state vector, each time the estimate, $\hat{e}(t)$, is obtained the model is updated. Thus if $\dot{z} = f(x, e, t)$ is the subvector of the complete state vector time derivative,

---

* $\delta(t - s)$ is the Dirac delta. Although w_e is not differentiable in the classical sense, u_e can be thought of as the "derivative" of w_e, and hence as an uncorrelated random variable with infinite covariance. A discussion of stochastic calculus is given in (5).

$\dot{x}(t)$, for which the errors are to be included, then $\hat{\dot{x}} = f(\hat{x}, \hat{e}, t)$ represents the updated portion of the model.

Note that in *(3.3.1)* or *(3.3.2)* the adaptive nature of the algorithm lies in the fact that the constant parameter vector, a, is estimated as well. Hence, if a reasonably accurate structure is assumed, subsequent estimation provides accurate values of a. There is a danger in using the forms as presented in *(3.3.1)* and *(3.3.2)*. As the estimation process is carried out, the a-vector converges to a constant value with the corresponding error covariance and hence gain terms approaching zero. Thus if the assumed structure is not an accurate approximation over the filter operating period, the eventual "hardening" of the structure as a takes on its constant value may, and usually will, produce divergence of the estimate. Since one either does not really know the error structure or else it is computationally too complex, it is usually a good idea to model the a-vector with a null structure, i.e., as a random parameter:

$$\dot{a}(t) = u_a, \quad E\{u_a\} = 0, \; E\{u_a(t)\, u_a^T(\tau)\} = Q_a(t)\, \delta(t - \tau) \qquad (3.3.3)$$

This keeps the error covariance from vanishing, and the gain will remain at some significant non-zero level such that the value of $a(t)$ is changed to allow the assumed structure to conform more accurately to the actual model errors.

Before discussing some specific examples of the structures used in the investigation, it will be advantageous to specifically define the error vector, e (dimensioned r), and parameter vector, a (dimensioned s), with regard to the previously mentioned forms. The differential approach is selected so that

$$\dot{e}(t) = g(e, a, t) + h(t)\, u_e(t)$$

$$\dot{a}(t) = u_a(t) \qquad (3.3.4)$$

We define the random variable

$$u^T(t) = [u_e(t) \vdots u_a(t)]$$

with statistics

$$E\{u(t)\} = 0, \quad E\{u(t)\, u^T(s)\} = Q(t)\, \delta(t - s)$$

For the cases where the algebraic form is to be employed, $e_0(t)$ is simply included in the parameter vector, a. The form (3.3.4) will be used as a general reference in the sequel, with the appropriate modifications for $e_0(t)$ being understood. Finally, the implementation is shown in Figures 1.1 and 1.2 with the modification that the state vector includes the components e and a; i.e.;

$$x(t) \text{ is redefined to be } \begin{bmatrix} x(t) \\ e(t) \\ a(t) \end{bmatrix} \tag{3.3.5}$$

where the $x(t)$ on the right is the original state vector.

A Simple Linear Structure. One of the simplest model error structures is a linear function of time,

$$e(t) = c_0 + c_1\, t \tag{3.3.6}$$

where

$$\dot{c}_0 = u_0$$

$$\dot{c}_1 = u_1$$

Hence,

$$a = \begin{bmatrix} c_0 \\ \cdots \\ c_1 \end{bmatrix}, \quad \dot{a}(t) = \begin{bmatrix} u_0 \\ \cdots \\ u_1 \end{bmatrix} = u_a(t) = u(t) \tag{3.3.7}$$

and

$$e(t) = [I \vdots t\ 1]\begin{bmatrix} c_0 \\ \cdots \\ c_1 \end{bmatrix} \qquad (3.3.8)$$

Note that without the presence of state noise, u_0 and u_1, the coefficient terms, c_0 and c_1, will take on nonvarying values, and the resulting "hardened" linear structure will probably not represent the model errors very well (unless, of course, the error actually *is* linear). The presence of the state noise keeps the gain from vanishing, and on the average (3.3.8) should approximate the model errors fairly well over a short interval.

Further discussion and applications of this model are given in Chapters 4 and 5.

The Ornstein-Uhlenbeck Stochastic Process. Another simple structure is that introduced by Ornstein and Uhlenbeck as a model to describe the velocity of a particle undergoing a Brownian motion (40, 516). This model has been used successfully by Ingram (32), Tapley and Ingram (41), and Tapley and Hagar (34), (35), in orbit determination studies. The Ornstein-Uhlenbeck process obeys a simple linear, first order differential equation (Langevin's equation) of the following type:

$$\dot{\varepsilon}(t) + \lceil \alpha \rfloor\, \varepsilon(t) = u_\varepsilon(t) \qquad (3.3.9)$$

where* α is a vector of constant parameters and $u_\varepsilon(t)$ is Gaussian white noise. The process is stationary, unbiased, and, by solution of (3.3.9), is exponentially correlated in time. Further discussions of the properties of this process are given in (5, 70-74) and (40, 516-524). For our purposes application to adaptive model error compensation gives

---

* The notation $\lceil \alpha \rfloor$ implies a diagonal matrix whose elements are the components of the vector, α.

$$\dot{e}(t) = \dot{\varepsilon}(t) = -\lceil \alpha \rfloor \varepsilon(t) + u_{\varepsilon}(t)$$

$$\dot{a}(t) = \dot{\alpha}(t) = u_{\alpha}(t)$$

$$u(t) = \begin{bmatrix} u_{\varepsilon} \\ \cdots \\ u_{\alpha} \end{bmatrix} \qquad (3.3.10)$$

Further discussion and applications of this model are given in Chapters 4 and 5.

Harmonically Bound Motion. If a particle undergoing Brownian motion is suddenly subjected to a displacement-proportional force, the equation of motion is simply that of a randomly forced harmonic oscillator (40, 524-525):

$$\ddot{\varepsilon}(t) + \lceil \alpha \rfloor \dot{\varepsilon}(t) + \lceil \beta \rfloor \varepsilon(t) = u_{\varepsilon}(t) \qquad (3.3.11)$$

where α and β are constant parameter vectors, and $u_{\varepsilon}(t)$ is Gaussian white noise. This process is stationary, unbiased, and has autocorrelation properties dependent upon the signs of $\lceil \alpha \rfloor$ and $\lceil \beta \rfloor$. This equation has been employed successfully as a model error compensation structure in low thrust space vehicle navigation by Tapley and Hagar, (34) and (35). Application of this structure to adaptive error compensation gives

$$\dot{e}(t) = \begin{bmatrix} \dot{\varepsilon}(t) \\ \cdots \\ \dot{\eta}(t) \end{bmatrix} = \begin{bmatrix} \eta(t) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ -\lceil \alpha \rfloor \eta(t) - \lceil \beta \rfloor \varepsilon(t) + u_{\varepsilon}(t) \end{bmatrix}$$

$$\dot{a}(t) = \begin{bmatrix} \dot{\alpha}(t) \\ \cdots \\ \dot{\beta}(t) \end{bmatrix} = \begin{bmatrix} u_{\alpha}(t) \\ \cdots \\ u_{\beta}(t) \end{bmatrix} = u_{a}(t)$$

$$u(t) = \begin{bmatrix} 0 \\ \cdots\cdots \\ u_{\varepsilon}(t) \\ \cdots\cdots \\ u_{\alpha}(t) \\ \cdots\cdots \\ u_{\beta}(t) \end{bmatrix} \qquad (3.3.12)$$

Further discussion and applications of this model are given in Chapters 4 and 5.

3.4 Linear Transformation of the Ornstein-Uhlenbeck Process

The model error forms presented in the last section were chosen in an apparently arbitrary fashion. In this section, a unifying derivation is presented from which each of the previous models can be obtained under appropriate assumptions. The general result is a representation shown to be equivalent to an exponential multiplying a Fourier series.

Consider the Ornstein-Uhlenbeck process of equation (3.3.9). Let $\varepsilon(t)$ be a $(2l + 1)$-vector related to the r-vector, $e(t)$, by a constant $r \times (2l + 1)$ matrix, C', whose values may be complex.

$$e(t) = C' \varepsilon(t) \tag{3.4.1}$$

Further, let $\varepsilon(t)$ satisfy the differential equation (3.3.9). For the j-th component of $e(t)$,

$$\begin{aligned} e_j(t) &= \sum_{m=1}^{2l+1} c'_{jm} \varepsilon_m(t) \\ &= \sum_{m=1}^{2l+1} c'_{jm} [\varepsilon_m(t_0) \exp(-\alpha_m t) + \int_{t_0}^{t_0+t} \exp(-\alpha_m s)\, w_m(s)\, ds] \\ &= \sum_{m=1}^{2l+1} c'_{jm} [\varepsilon_m(t_0) \exp(-\alpha_m t) + w'_w] \end{aligned} \tag{3.4.2}$$

where the α_m are also allowed to be complex. Taking the expected value, the second term in brackets vanishes. By absorbing the $\varepsilon(t_0)$ into the c'_{jm}, one then obtains

$$E\{e_j(t)\} \equiv e_j^*(t) = \sum_{m=1}^{2l+1} c_{jm} \exp(-\alpha_m t) \tag{3.4.3}$$

with $c_{jm} = c'_{jm} \varepsilon_m(0)$. This is the general form of the transformation. With c_{jm} and α_m complex, the only restriction is that $e_j^*(t)$ is usually required to be real in order to represent real model errors. If α_m is further restricted to be of the form

$$\alpha_m - \rho_j + i\, m\, \omega_j \qquad (3.4.4)$$

where $i = \sqrt{-1}$, and if the limits on the summation are changed to run from $-l$ to $+l$, (3.4.3) becomes

$$e_j^*(t) = \exp(-\rho_j t) \sum_{m=-l}^{l} c_{jm} \exp(-i\, m\, \omega_j t) \qquad (3.4.5)$$

With c_{jm} of the form

$$c_{j0} = \frac{s_{j0}}{2}$$

$$c_{jm} = \frac{s_{jm} - i\, r_{jm}}{2}, \quad m > 0$$

$$c_{jm} = \frac{s_{jm} + i\, r_{jm}}{2}, \quad m < 0$$

then equation (3.4.5) can be written as

$$e_j^*(t) = \exp(-\rho_j t) \sum_{m=0}^{l} \{s_{jm} \cos(m\, \omega_j t) + r_{jm} \sin(m\, \omega_j t)\} \qquad (3.4.6)$$

where use has been made of the definitions

$$\cos\phi = [\exp(i\,\phi) + \exp(-i\,\phi)]/2, \quad \sin\phi = [\exp(i\,\phi) - \exp(-i\,\phi)]/2i$$

Thus under the foregoing restrictions, the process reduces to an exponential multiplying a truncated Fourier series. Some special cases are now considered which yield the models of section 3.3.

Case 1. Let $l = 1$, $\rho_j = 0$, and using the series expansion for $\sin$ and $\cos$ through the first order, (3.4.6) becomes

$$e_j^*(t) = s_{j1} + r_{j1}\, \omega_j\, t \qquad (3.4.9)$$

which becomes the linear form (3.3.6) upon making the identifications

$$c_0 = s_{j1}, \quad c_0 = \begin{bmatrix} c_{10} \\ \vdots \\ c_{r0} \end{bmatrix}, \qquad c_{j1} = r_{j1}\, \omega_j, \quad c_1 = \begin{bmatrix} c_{11} \\ \vdots \\ c_{r1} \end{bmatrix}$$

Case 2. Let $l = 0$, but retain the exponential form. Then (3.4.6) gives

$$e_j^*(t) = \varepsilon_{j0} \exp(-\rho_j t) \tag{3.4.8}$$

which is the solution to the differential equation

$$\dot{e}_j^*(t) = -\rho_j e_j^*(t)$$

This is obviously of the form of the mean of (3.3.9) with

$$\alpha_j = \rho_j$$

Case 3. Let $l = 1$, and set $\varepsilon_{j0} = 0$. Then (3.4.6) becomes

$$e_j^*(t) = \exp(-\rho_j t)[\varepsilon_{j1} \cos \omega_j t + r_{j1} \sin \omega_j t] \tag{3.4.9}$$

which is the solution to the harmonic oscillator

$$\ddot{e}_j^*(t) + 2\rho_j \dot{e}_j^*(t) + (\rho_j^2 - \omega_j^2) e_j^*(t) = 0$$

Making the identifications

$$\alpha_j = 2\rho_j, \quad \beta_j = \rho_j^2 - \omega_j^2$$

the result is of the form of the mean of (3.3.11).

It should be apparent that more complex structures can be obtained easily from (3.4.6). For example, without restricting ε_{j0} to be zero, the resulting form of (3.4.6) with $l = 1$ satisfies the nonhomogeneous differential equation,

$$\ddot{e}_j^*(t) + \rho_j \dot{e}_j^*(t) + \omega_j^2 e^*(t) + \omega_j^2 \varepsilon_{j0} \exp(-\rho_j t) = 0$$

3.5 Structural Adaptation

With the general form (3.4.6) established as being capable of representing a number of different functions, it is possible, at least in theory, to adapt structurally to a close approximation of the dynamic model errors. By defining the vector of model errors and parameters,

$$z(t) \equiv \begin{bmatrix} e_1(t) \\ \vdots \\ e_m(t) \\ \vdots \\ \rho_1(t) \\ \vdots \\ \rho_m(t) \\ \omega_1(t) \\ \vdots \\ \omega_m(t) \end{bmatrix}$$

and selecting m to be large enough (based on the degree of sophistication of the *a priori* dynamic model), the filtering equations (Algorithms I or II, Chapter 1) will produce values for the ρ_i and ω_i which will result in an approximation, e_j, to the true structure of the model errors. The obvious disadvantage is in the implementation: a large m produces a very significant computational load. Thus the flexibility for structural adaptation, as offered by a form such as *(3.4.6)*, is partially offset by the need to exercise parexic judgment in selecting a value of m.

3.6 A Sequential State Noise Covariance Estimator

In this section we return to the problem of estimating the state noise covariance matrix. A new approach is taken, although the technique is simply the formulation of a Kalman filter algorithm applied to a vector whose elements are those of the state noise covariance matrix, Q. With this objective in mind, once we have obtained the analogous state-observation and state dynamics linear relationships, the appropriate estimation equations can be written down immediately. The resulting method will be shown to yield a minimum variance estimate of the state noise covariance matrix elements, subject to the additional restriction that Q be non-negative definite.

We begin by establishing the square of the kth predicted residual as the required observation. From section 3.1, the predicted residual is defined as

$$\begin{aligned}\bar{r}_k &= y_k - H_k \bar{x}_k \\ &= y_k - H_k \Phi_{k,k-1} \hat{x}_{k-1} \\ &= H_k \Phi_{k,k-1} \tilde{x}_{k-1} + H_k \Gamma_{k,k-1} w_{k-1} + v_k \end{aligned} \tag{3.6.1}$$

Continuing with the previous assumption that we are dealing with scalar observations, we then form

$$E\{\bar{r}_k \bar{r}_k^T\} = E\{\bar{r}_k^2\} = H_k \bar{P}_k H_k^T + R_k \tag{3.6.2}$$

and define the error, θ, as

$$\theta_k = \bar{r}_k^2 - E\{\bar{r}_k^2\} \tag{3.6.3}$$

where $E\{\theta\} = 0$. Then

$$\bar{r}_k^2 = H_k \bar{P}_k H_k^T + R_k + \theta_k \tag{3.6.4}$$

Further, assume that R_k is exactly known and define

$$\bar{D}_k = \bar{r}_k^2 - R_k \tag{3.6.5}$$

$$V_k = H_k \bar{P}_k H_k^T \tag{3.6.6}$$

so that

$$\bar{D}_k = V_k + \theta_k \tag{3.6.7}$$

Now V_k is a function of Q_{k-1}, since

$$\bar{P}_k = \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^T + \Gamma_{k,k-1} Q_{k-1} \Gamma_{k,k-1}^T \tag{3.6.8}$$

and it is Q_{k-1} that we are trying to estimate. If a Kalman filtering algorithm is to be developed, then a linear equation relating Q_{k-1} to $\bar{D}$ must be obtained. To this end, the elements of Q are first placed in a vector, q,

such that the elements are stored row-wise. Further, use is made of the fact that Q is symmetric so that only the upper triangular portion is stored. Thus,

$$q^T = (q_{11} \cdots q_{1m}\, q_{22} \cdots q_{2m}\, q_{33} \cdots q_{m-1,m}\, q_{mm}) \tag{3.6.9}$$

where $m \times m$ is the dimension of Q, and $q_{k-1} = q(t_{k-1})$ is therefore the $m(m+1)/2 \times 1$ state vector to be estimated. Proceeding formally, consider a Taylor series expansion for V_k expanded about the value, $q_{k-1} = 0$, and truncated to the linear term:

$$V_k(q_{k-1}) = V_k \Big|_{q_{k-1}=0} + \frac{\partial V_k}{\partial q_{k-1}} \Big|_{q_{k-1}=0} q_{k-1} \tag{3.6.10}$$

$V_k(0)$ is simply (3.6.6) with $Q_{k-1} = 0$. In analogy with the linearized observation state equation (1.2.23), we use (3.6.10) and write (3.6.7) as

$$\overline{\Delta}_k = J_k\, q_{k-1} + \theta_k \tag{3.6.11}$$

where

$$\overline{\Delta}_k \equiv \overline{D}_k - V_k(0) = \overline{r}_k^{\,2} - H_k\, \overline{P}_k\, H_k^T \Big|_{q_{k-1}=0} - R_k$$

$$J_k = \frac{\partial V_k}{\partial q_{k-1}} \Big|_{q_{k-1}=0} \tag{3.6.12}$$

The final task in developing the components of the observation equation is to determine the elements of J_k. Since $\overline{P}_k$, and hence V_k, is linear in q_{k-1}, equation (3.6.10) is exact; we write (3.6.10)

$$V_k(q_{k-1}) = V_k(0) + \delta V_k(0) \tag{3.6.13}$$

The expansion of V_k with respect to Q_{k-1} yields

$$\delta V_k = \delta(H_k \bar{P}_k H_k^T) = H_k \Gamma_{k,k-1} \delta Q_{k-1} \Gamma_{k,k-1} H_k^T$$

$$= H_k \Gamma_{k,k-1} Q_{k-1} \Gamma_{k,k-1}^T H_k^T \qquad (3.6.14)$$

where $\delta Q_{k-1} = Q_{k-1} - Q_{k-1}^* = Q_{k-1}$ since the expansion is about the nominal value $Q_{k-1}^* = 0$. Now the key factor in the development hinges on the ability to express the elements of $H\,Q\,H^T$ in the form

$$J\,q^T \qquad (3.6.15)$$

where J is composed of the elements of H and Γ. To this end we digress for a moment and prove the following:

Theorem I. Given the matrices A, B, C and X of dimensions $l \times m$, $m \times n$, $nl \times m^2$, and $m \times m$ respectively, then the $l \times n$ matrix, AXB, can be written in the vector form, Cx, where $x^T = (x_{11}\, x_{12} \cdots x_{1m}\, x_{22}\, x_{23} \cdots x_{2m} \cdots x_{mm})$, an m^2-vector, and $C = \{c_{kp}\}$ with the c_{ij} being given by

$$c_{ij} = a_{\lfloor (i-1)/n+1 \rfloor, \lfloor (j-1)/m+1 \rfloor}\, b_{[mod(j-1,m)\,+\,1)],\,[mod(i-1,n)\,+\,1]}$$

where

$\lfloor s \rfloor$ = *integer part of* s

and

$$mod(s,\ t) = s - \left\lfloor \frac{s}{t} \right\rfloor t$$

Proof. Writing AXB in expanded form results in the following sequence:

$$AXB = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{l1} & \cdots & a_{m} \end{bmatrix} \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mm} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{l1} & \cdots & a_{lm} \end{bmatrix} \begin{bmatrix} \sum_j^m x_{1j}\, b_{j1} & \cdots & \sum_j^m x_{1j}\, b_{jn} \\ \vdots & & \vdots \\ \sum_j^m x_{mj}\, b_{j1} & \cdots & \sum_j^m x_{mj}\, b_{jn} \end{bmatrix}$$

$$AXB = \begin{bmatrix} \sum_i^m a_{1i} \sum_j^m x_{ij} b_{j1} & \cdots & \sum_i^m a_{1i} \sum_j^m x_{ij} b_{jn} \\ \vdots & & \vdots \\ \sum_i^m a_{li} \sum_j^m x_{ij} b_{j1} & \cdots & \sum_i^m a_{li} \sum_j^m x_{ij} b_{jn} \end{bmatrix} \tag{3.6.16}$$

The general element of AXB is further expanded as

$$\sum_i^m \sum_j^m a_{ki} b_{jp} x_{ij} = a_{k1} b_{1p} x_{11} + a_{k1} b_{2p} x_{12} + \cdots a_{k1} b_{mp} x_{1m}$$
$$+ a_{k2} b_{1p} x_{21} + a_{k2} b_{2p} x_{22} + \cdots + a_{k2} b_{mp} x_{2p}$$
$$+ \cdots a_{km} b_{mp} x_{mm}, \quad k = 1,m, \; p = 1,m$$

In vector form this becomes

$$\sum_i^m \sum_j^m a_{ki} b_{jp} x_{ij} = (a_{k1} b_{1p} \; a_{k1} b_{2p} \cdots a_{km} b_{mp}) \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{mm} \end{bmatrix} \tag{3.6.17}$$

Arranging the elements of (3.6.16) into a column vector and using (3.6.17) results in

$$\begin{bmatrix} \sum_i^m \sum_j^m a_{1i} b_{j1} x_{ij} \\ \vdots \\ \sum_i^m \sum_j^m a_{1i} b_{jn} x_{ij} \\ \vdots \\ \sum_i^m \sum_j^m a_{li} b_{j1} x_{ij} \\ \vdots \\ \sum_i^m \sum_j^m a_{li} b_{jn} x_{ij} \end{bmatrix} = \begin{bmatrix} a_{11} b_{11} & \cdots & a_{1m} b_{m1} \\ \vdots & & \vdots \\ a_{11} b_{1n} & \cdots & a_{1m} b_{mn} \\ \vdots & & \vdots \\ a_{l1} b_{11} & \cdots & a_{lm} b_{m1} \\ \vdots & & \vdots \\ a_{l1} b_{1n} & \cdots & a_{lm} b_{mn} \end{bmatrix} \begin{bmatrix} x_{11} \\ \vdots \\ x_{1m} \\ x_{21} \\ \vdots \\ x_{2m} \\ \vdots \\ x_{mm} \end{bmatrix} \tag{3.6.18}$$

Defining the elements in the $nl \times m^2$ matrix on the right as the elements of C completes the proof.

In view of this theorem we also have the following:

Corollary. Given the conditions of the Theorem I above, let X and AXB be symmetric. Then AXB may be placed in the vector form Cx where $x^T = (x_{11} \ldots x_{1m}\ x_{22} \ldots x_{2m} \ldots x_{mm})$ is a vector of $m(m+1)/2$ elements and C is an $ln(ln+1)/2 \times m(m+1)/2$ matrix with elements

$$c_{ij} = \begin{cases} a_{pq}\, b_{rs} \ , & \text{for } q = r \ , \\ a_{pq}\, b_{rs} + a_{pr}\, b_{qs} \ , & \text{for } q \neq r \end{cases} \qquad (3.6.19)$$

and where

$$p = \lfloor (i+t-1)/n + 1 \rfloor$$

$$q = \lfloor (j+u-1)/m + 1 \rfloor$$

$$r = [\mathrm{mod}(j+u-1,\ m) + 1]$$

$$s = [\mathrm{mod}(i+t-1,\ n) + 1]$$

$$t = \lfloor i/n \rfloor \quad \text{and} \quad u = \lfloor j/m \rfloor$$

The equations for the indeces of the terms in Theorem I and the corollary follow by induction.

Since X is symmetric, $x_{ij} = x_{ji}$, $i \neq j$. Then the redundant elements of x may be eliminated so that $x^T = (x_{11} \ldots x_{1m}\ x_{22} \ldots x_{2m}\ x_{33} \ldots x_{3m} \ldots x_{mm})$ is $m(m+1)/2$. In Cx the elements of C corresponding to x_{ij}, for $i > j$, are simply added to those corresponding to x_{ij}, for $i < j$. Further, the symmetry of AXB yields

$$\sum_{i}^{m} \sum_{j}^{m} a_{ki}\, x_{ij}\, b_{jp} = \sum_{i}^{m} \sum_{j}^{m} a_{pi}\, x_{ij}\, b_{jk}$$

hence the duplicated rows of C need not be included. As a result, the elements of C are given by (3.6.19).

In light of this corollary, the validity of *(3.6.15)* is established. A simple example will serve to illustrate the application. Consider a 2×2 Q-matrix, with H^T as a 3-vector. Then

$$Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$$

$$H = [h_1 \; h_2 \; h_3]$$

$$\Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \\ \gamma_{31} & \gamma_{32} \end{bmatrix}$$

and

$$H\Gamma = [h_1 \; h_2 \; h_3] \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \\ \gamma_{31} & \gamma_{32} \end{bmatrix} = \begin{bmatrix} \sum_1^3 h_i \gamma_{i1} & \sum_1^3 h_i \gamma_{i2} \end{bmatrix} \qquad (3.6.20)$$

Making the identifications

$$A = B^T = H\Gamma, \quad X = Q$$

then application of the corollary to Theorem I yields

$$x^T \equiv q^T = (q_{11} \; q_{12} \; q_{22})$$

and

$$C \equiv J = \left[\left(\sum_1^3 h_i \gamma_{i1} \right)^2 \; 2 \left(\sum_1^3 h_i \gamma_{i1} \right) \left(\sum_1^3 h_i \gamma_{i2} \right) \; \left(\sum_1^3 h_i \gamma_{i2} \right)^2 \right]$$

Note that in cases where J has only one row (as in this example), then we can write the *equality*

$$H \Gamma Q \Gamma^T H^T = J q \qquad (3.6.21)$$

This example corresponds to the case of scalar observations. In line with our previous developments we continue with this assumption, although it is not a restriction.

With the validity of *(3.6.11)* established and its elements determined, it is important to consider the variance of $\bar{\Delta}_k$. This is necessary as the inverse of this variance and some of its elements appear in the filter gain.

Defining

$$\tilde{\Delta}_k = J_k \tilde{\tilde{q}}_{k-1} + \theta_k = J_k(q_{k-1} - \bar{q}_{k-1}) + \theta_k \qquad (3.6.22)$$

and

$$G_k \equiv E\{\tilde{\Delta}_k \tilde{\Delta}_k^T\} = E\{\tilde{\Delta}_k^2\}$$

then

$$G_k = J_k E\{\tilde{\tilde{q}}_{k-1} \tilde{\tilde{q}}_{k-1}^T\} J_k^T + J_k E\{\tilde{\tilde{q}}_{k-1} \theta_k\} + E\{\theta_k \tilde{\tilde{q}}_{k-1}^T\} J_k^T + E\{\theta_k^2\} \qquad (3.6.23)$$

Letting

$$\bar{S}_k = E\{\tilde{\tilde{q}}_{k-1} \tilde{\tilde{q}}_{k-1}^T\}$$

$$\bar{C} = E\{\tilde{\tilde{q}}_{k-1} \theta\}$$

and

$$T_k = E\{\theta_k^2\} \qquad (3.6.24)$$

then *(3.6.23)* becomes

$$G = J \bar{S} J^T + 2 J \bar{C} + T \qquad (3.6.25)$$

In the ordinary Kalman filter the observation error is assumed to be uncorrelated with the *a priori* state estimate. Here this case is not generally true; both θ and $\tilde{\tilde{q}}$ are functions of u, and this correlation is reflected in C. However, it is difficult at best to determine appropriate values for the elements of C. Thus, in view of this lack of knowledge, a viable alternative is simple to assume $C = 0$ so that *(3.6.24)* becomes

$$G = J \bar{S} J^T + T$$

This expression is directly analogous to the observation residual variance of the ordinary Kalman filter.

We now lack only a "dynamics" relationship governing the prediction of q. Assume that q_{k-1} is linearly related to q_{k-2} by an appropriate state transition matrix, $\Psi(t_{k-1}, t_{k-2})$:

$$q_{k-1} = \Psi_{k-1,k-2}\, q_{k-2} \tag{3.6.26}$$

If the process generating the elements of state noise vector, u, is wide-sense stationary, then the state noise covariance matrix is constant. Making this assumption we have

$$q_{k-1} = q_{k-2}$$

and hence Ψ is the identity matrix, I. Such an assumption may not be overly restrictive since usually the dynamic model governing x can be determined such that u is at least bounded over the interval of interest, $t_0 \leq t \leq t_f$.

We are now in a position to write down the estimation equations. Making use of the definitions (3.6.24) and employing the assumption $C = 0$, the appropriate relationships may be written directly as

$$\bar{S}_k = S_{k-1}, \quad \bar{q}_{k-1} = \hat{q}_{k-2} \tag{3.6.27a}$$

$$M_k = \bar{S}_k J_k^T / (J_k \bar{S}_k J_k^T + T_k) \tag{3.6.27b}$$

$$\hat{q}_{k-1} = \bar{q}_{k-1} + M_k(\Delta_k - J_k \bar{q}_{k-1}) \tag{3.6.27c}$$

$$S_k = (I - M_k J_k) \bar{S}_k (I - J_k^T M_k^T) + M_k T_k M_k^T \tag{3.6.27d}$$

where M_k is the filter gain, and

$$S_k = E\{(q_{k-1} - \hat{q}_{k-1})(q_{k-1} - \hat{q}_{k-1})^T\} \tag{3.6.28}$$

There are additional restrictions which must be set forth before the algorithm may be used. Recall that Q must be non-negative definite. However, it is

not unlikely that (3.6.27a) may produce an estimate $\hat{q}$ which does not conform to this requirement, particularly during initial (transient) operation of the filter. In order to accommodate the non-negative definite constraint we call upon the properties of the covariance, Q. First, recall that for the diagonal elements, q_{ii},

$$q_{ii} \geq 0 \tag{3.6.29}$$

Also, the q_{ij}, $i \neq j$, may be written in terms of the correlation coefficient, ρ_{ij}, as

$$q_{ij} = \rho_{ij} \sqrt{q_{ii}\, q_{jj}} \tag{3.6.30}$$

where

$$-1 \leq \rho_{ij} \leq 1$$

Now since (3.6.29) and (3.6.30) are inherent properties of a covariance matrix, and since any covariance matrix must be non-negative definite, (3.6.29) and (3.6.30) may be used to insure the restriction on Q is met. Thus if any diagonal element is estimated to be negative, it is then set to zero. Correspondingly the appropriate gain element is recomputed to agree with this modification:

$$0 = \bar{q}_{ii} + m(\bar{\Delta} - J\,\bar{q})$$

or

$$m = -\bar{q}_{ii}/(\bar{\Delta} - J\,\bar{q}) \tag{3.6.31}$$

where m is the corresponding element of M_k. After performing this test, if

$$|\hat{q}_{ij}| > \sqrt{\hat{q}^m_{ii}\, \hat{q}^m_{jj}} \tag{3.6.32}$$

then set

$$\hat{q}_{ij} = max\{ -\sqrt{\hat{q}_{ii}\, \hat{q}_{jj}}\,,\ min[\sqrt{\hat{q}_{ii}\, \hat{q}_{jj}}\,,\ q_{ij}{}^m]\} \tag{3.6.33}$$

where the superscript, m, means the value computed by *(3.6.25)*, and where the *max* and *min* functions mean take the respective maximum or minimum values of the arguments. Further, the corresponding gain term is modified as

$$m = (\hat{q}_{ij} - \bar{q}_{ij})/(\bar{\Delta} - J\,\bar{q}) \qquad (3.6.34)$$

The foregoing restrictions do not destroy the unbiased property of the estimate, i.e., $E\{\hat{q}\} = q$. Further, since the Kalman filter normally produces a minimum variance estimate, the restrictions on q result in a constrained minimum variance estimate. Of course, this is true only within the validity of the assumption $E\{q\theta\} = 0$.

No mention has yet been made as to the choice of the residual error variance, T. Since it is somewhat difficult to select this value, it may be estimated also. One way is to use a method proposed by Tapley and Born (43). The technique is simple and straightforward, and is the average of the *a posteriori* residuals given by the following recursive equation:

$$T_k = (1 - \tfrac{1}{k})\,T_{k-1} + \tfrac{1}{k}\,(\bar{\Delta}_k - J_k\,\hat{q}_{k-1})^2 \qquad (3.6.35)$$

An initial value which may be used is $T_0 = \bar{\Delta}_0^2$, since if $J\,q = 0$,

$$\bar{\Delta}^2 = \theta^2 \simeq T$$

Since *(3.6.35)* is the average of all the residuals up through t_k, as k becomes larger, each new residual has less effect in determining T_k. This is acceptable if θ is wide-sense stationary, in which case T is constant. If this is not the case, then a modification which offsets this effect is to use only the last N residuals. Thus for $k \leq N$, equation *(3.6.35)* is employed. At each t_k the kth residual is saved. Then for $k > N$ use

$$T_k = T_{k-1} + \tfrac{1}{N}\,(\bar{\Delta}_k - J_k\,\hat{q}_{k-1})^2 - \tfrac{1}{N}\,(\bar{\Delta}_{k-N} - J_{k-N}\,\hat{q}_{k-N-1})^2 \qquad (3.6.36)$$

instead of *(3.6.35)*, and continue to save the most current N residuals.

The implementation of the sequential state noise covariance estimation algorithm is diagrammed in Figure 3.5.

Although we have been concerned here with estimating only the elements of Q, it should be obvious that the algorithm may be extended to include the simultaneous estimation of the observation error variance R. In this case, an augmented state vector, z, is defined as

$$z = (q \; R)^T$$

Then $\bar{\Delta}$ and J are redefined as

$$\bar{\Delta} = \bar{r}^2 - H \bar{P} H^T = J z + 0$$

and

$$J = (G \; 1) \qquad (3.6.37)$$

The corresponding estimation equations are of the same form, although $\bar{P}$ is replaced by $\bar{r}^2$, and z replaces q.

It was indicated earlier that the assumption of scalar observations is not a restriction. Consider the case where $H\Gamma$ is $l \times m$. Then $H\Gamma Q \Gamma^T H^T$ is $l \times l$. Correspondingly,

$$\bar{r} \, \bar{r}^T = H \Phi P \Phi^T H^T + H \Gamma Q \Gamma^T H^T + R + 0 \qquad (3.6.38)$$

is also $l \times l$. Since the terms of (3.6.38) are symmetric, the corollary to Theorem I applies, and we can write

$$*\bar{r}^2 = *(H \Phi P \Phi^T H^T) + J q + *R + *\theta \qquad (3.6.39)$$

where each of the terms is $l(l + 1)/2 \times 1$. For instance, if $l = 2$, $*\bar{r}^2 = (\bar{r}_1^2 \; \bar{r}_1\bar{r}_2 \; \bar{r}_2^2)^T$ with similar arrangements for $*(H\Phi P\Phi^T H^T)$, $*R$, and $*\theta$. Thus,

$$*(H \Phi P \Phi^T H^T) = *V(0) = \begin{bmatrix} V_{11} \\ V_{12} \\ V_{22} \end{bmatrix}_{q=0}$$

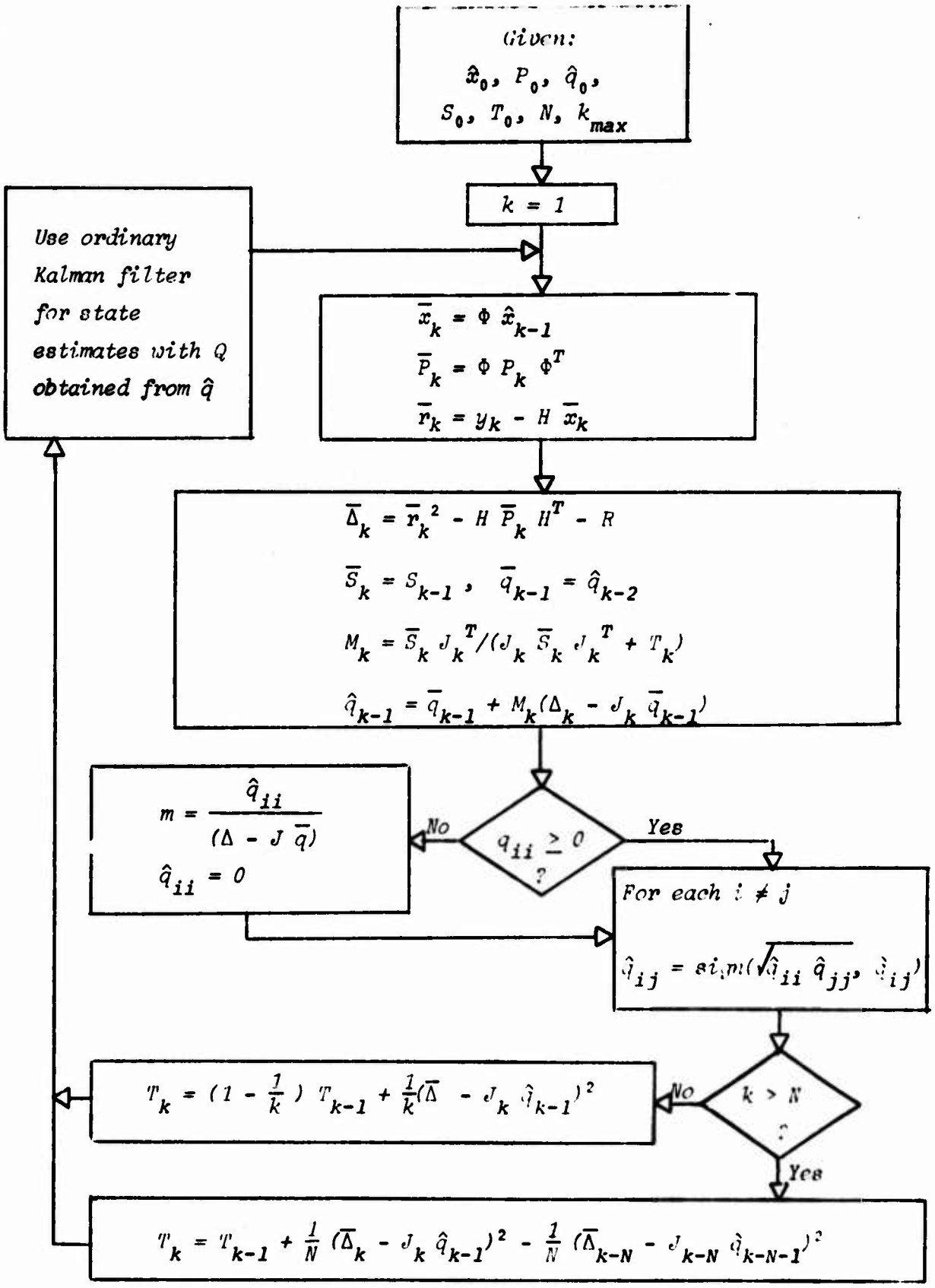


Figure 3.5 State Noise Covariance Sequential Estimator

and

$$^{*}R = \begin{bmatrix} R_{11} \\ R_{12} \\ R_{22} \end{bmatrix}$$

and

$$^{*}\theta = \begin{bmatrix} \theta_{11} \\ \theta_{12} \\ \theta_{22} \end{bmatrix} \qquad (3.6.40)$$

The rest of the algorithm is obtained in a straightforward manner, similar to that for Algorithm I (Chapter 1), subject to the non-negative definite restrictions for $\hat{q}$.

Finally we note that the estimator developed above suffers certain disadvantages. One of the more serious is the assumption of a specific model for Q. This may be reasonable for a portion of the filter operating range. However, if the state noise process, u, changes its behavior drastically after steady state conditions have been essentially reached, the filter has no power to adapt to the corresponding new value of Q. Recall that this is due to the fact that the associated error covariance practically vanishes, and, similarly, so does the corresponding gain. This situation is exactly analogous to the original dynamic model error problem, and is the most serious disadvantage of this Q-estimator. However, this being the case, we can apply any of the previously developed compensation algorithms, including another state noise covariance sequential estimator.

For example, an obvious method is to include a state noise fourth moment matrix*, W. In this way the prediction of S_k given by *(3.6.27a)* is

---

* By using another filter to estimate W, this approach could be continued *ad infinitum*, becoming computationally very burdensome very quickly. The author's opinion is that anything much more complex than *(3.6.41)* is seldom practical.

modified to

$$\bar{S}_k = S_{k-1} + W_{k-1} \qquad (3.6.41)$$

3.7 Summary

This chapter has presented a selection of adaptive optimal and suboptimal algorithms to compensate for model errors. In all cases the emphasis has been on simplicity and generality. Jazwinski's algorithm for estimating the state noise covariance matrix has been shown to be optimal based on a single predicted residual. By using the mean of a sample of squared predicted residuals, an element of smoothing is introduced, but also at the expense of introducing suboptimality.

The basis of Jazwinski's algorithm has also been shown to be applicable to a number of the algorithms of Chapter 2, allowing the suboptimal parameters to be adaptively determined.

In section 3.3 the approach (different from all the previous ones) of attempting to estimate the actual model errors has been taken. This adaptive method has been shown to require the assumption of a particular functional form to represent the structure of the model error. Several examples have been presented, and a general functional form developed to allow a measure of structural adaptation. This has been in the form of a generalized Ornstein-Uhlenbeck process, capable of structurally approximating (having the approximate functional form of) a number of different functions.

A Kalman filtering algorithm for sequentially estimating the state noise covariance matrix, Q, has been developed in section 3.6. The algorithm is somewhat more complex than those previously presented, but has been shown to provide a minimum variance estimate of Q, subject to the constraint that Q be non-negative definite.

Chapter 4

APPLICATIONS: RENDEZVOUS MISSION

4.1 Introduction

In this chapter the algorithms presented in the foregoing chapters are applied to the rendezvous problem formulated in Section 1.4. The interest here is to gain familiarity with the operation of the various estimation algorithms. Through the development of explicit equations, and plots of their behavior, qualitative insight may be gained into the performance of the algorithms.

4.2 Age-Weighting of Data

The standard age-weighting algorithm is easily applied to the rendezvous problem discussed in Section 1.4. Modifying the estimation equations *(1.4.6)* to conform to the age-weighting algorithm, equations *(2.1.1)*, result in the following sequence:

$$\bar{\dot{\rho}}_k = \hat{\dot{\rho}}_{k-1}$$

$$\bar{P}_k = P_{k-1}$$

$$K_k = \bar{P}_k[\bar{P}_k + R/s]^{-1}$$

$$\hat{\dot{\rho}}_k = \bar{\dot{\rho}} + K[y - \bar{\dot{\rho}}]$$

$$P_k = s[1 - K]\,\bar{P} \qquad (4.2.1)$$

where it is tacitly assumed that R is constant. Starting with $P(t_0) = P_0$, the gain and covariance recursions above may be applied to arrive at the

steady state gain and covariance values*. Thus,

$k = 1$:

$$\bar{P}_1 = P_0$$

$$K_1 = \bar{P}_1[\bar{P}_1 + R/s]^{-1}$$

$$= \frac{P_0}{(P_0 + R/s)}$$

$$P_1 = \left[1 - \left[\frac{P_0}{(P_0 + R/s)}\right]\right] P_0$$

$$= \frac{R\,P_0}{(P_0 + R/s)}$$

$k = 2$:

$$\bar{P}_2 = P_1 = \frac{P_0\,R}{(P_0 + R/s)}$$

$$K_2 = \bar{P}_2[\bar{P}_2 + R/s]^{-1}$$

$$= \frac{P_0\,R/(P_0 + R/s)}{\left[\frac{P_0\,R}{(R_0 + R/s)} + R/s\right]}$$

$$= P_0/\left[P_0 + \frac{P_0}{s} + \frac{R}{s^2}\right]$$

$$P_2 = \left[\frac{P_0/s + R/s^2}{P_0(1 + 1/s) + R/s^2}\right] \frac{P_0\,R}{(P_0 + R/s)}$$

$$= P_0\,R/[P_0(1 + 1/s) + R/s^2]$$

---

* The gain and covariance update equations could be used to solve directly for the steady state conditions. However, using the recursion equations to compute values recursively and then inductively determining the steady state conditions provide insight into the operation of the algorithm.

$k = 3$:

$$\bar{P}_3 = P_2 = \frac{P_0 R}{P_0(1 + 1/s) + R/s^2}$$

$$K_3 = \bar{P}_3[\bar{P}_3 + R/s]^{-1}$$

$$= \frac{P_0 R/[P_0(1 + 1/s) + R/s^2]}{P_0 R/[P_0(1 + 1/s) + R/s^3] + R/s}$$

$$= \frac{P_0}{P_0 + P_0(1/s + 1/s^2) + R/s^3} = \frac{P_0}{P_0(1 + 1/s + 1/s^2) + R/s^3}$$

$$P_3 = P_0 R/[P_0(1 + 1/s + 1/s^2) + R/s^3]$$

Continuing the recursion, one obtains for $k = m$:

$$K_m = \frac{P_0}{P_0 \sum_{j=0}^{m-1} 1/s^j + R/s^m}$$

$$P_m = \frac{P_0 R}{P_0 \sum_{j=0}^{m-1} 1/s^j + R/s^m} \qquad (4.2.2)$$

The summation in K_m and P_m is simply a geometric series with sum, as $m \to \infty$,

$$\sum_{j=0}^{\infty} 1/s^j = \frac{1}{1 - 1/s} \qquad (4.2.3)$$

Thus, as $m \to \infty$, K_m and P_m become

$$K_\infty = \frac{P_0}{P_0\left(\frac{s}{s - 1}\right)} = \frac{s - 1}{s} \qquad (4.2.4)$$

where $\lim_{m\to\infty} R/s^m = 0$, since $s > 1$, and

$$P_\infty = \frac{s - 1}{s} R \qquad (4.2.5)$$

For very large k, the estimate of $\dot{\rho}$, including the first order model error effects, is found to be approximately

$$\hat{\dot{\rho}}_k \approx \bar{\dot{\rho}}_k + \frac{s-1}{s}\left(\rho_0\gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k - \bar{\dot{\rho}}_k + v_k\right)$$

or

$$\hat{\dot{\rho}}_k \approx \frac{1}{s}\bar{\dot{\rho}}_k + \frac{s-1}{s}\left(\rho_0\gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k + v_k\right) \qquad (4.3.6)$$

Using (1.4.25), the corresponding error is given approximately by the following expression

$$\tilde{\dot{\rho}}_k \approx \frac{1}{s}\left(\rho_0\gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k\right) - \frac{1}{s}\bar{\dot{\rho}} - \frac{s-1}{s}v_k \qquad (4.3.7)$$

For large s, the effect of the unbounded error term (first term on the right) is small. Correspondingly, the random effects are more predominant, and in the limit as $s \to \infty$ the error is due strictly to measurement noise. Because the time correlated term completely dominates the noise term when k is large, continued filter operation is possible only when s is very large. In fact, best performance results here when s is very large since then all the information is derived from the most current observations. As expected (section 2.1) the error covariance in this case becomes, by (4.2.5), R.

4.3 Schmidt Suboptimal Filter: Gain Scaling

In section 2.2 it was indicated that the use of a constant value of b can lead to certain difficulties. This is exemplified here when one attempts to find the steady state gain and error variance. The error covariance recursion is used to solve for the steady state variance directly:

$$P_m = \left[1 - (2b - b^2)\frac{\bar{P}_m}{\bar{P}_m + R}\right]\bar{P}_m$$

In steady state, $P_m = P_\infty = P$.

$$P = [1 - (2b - b^2) \frac{P}{(P + R)}] P$$

$$0 = -(2b - b^2) \frac{P}{P + R} \qquad (4.3.8)$$

This implies either $P_\infty = 0$ or $b = 2$. Thus for $1 < b < 2$, it simply takes longer for the steady state condition of $P = 0$ (and hence $K = 0$) to be reached. For $b = 2$, $P = P_0$. In this case, the estimate eventually diverges.

For the method where (2.2.12) and (2.2.13) are employed, the steady state variance, P_∞, is found not to vanish, since in this case P_∞ is found to be

$$P_\infty = \frac{\alpha R}{2 - \alpha} \qquad (4.3.9)$$

The corresponding gain is simply

$$K_\infty = \alpha \qquad (4.3.10)$$

As $\alpha \to 1$, only the most recent observations contribute information. Consequently, $K_\infty \to 1$ and $P_\infty \to R$.

The corresponding estimate and error for large k are determined as

$$\hat{\dot{\rho}}_k \simeq \bar{\dot{\rho}}_k + \frac{\varepsilon}{R} (\rho_0 \gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k - \bar{\dot{\rho}}_k + v_k)$$

$$\sim (1 - \frac{\varepsilon}{R}) \bar{\dot{\rho}}_k + \frac{\varepsilon}{R} (\rho_0 \gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k + v_k) \qquad (4.3.11)$$

and

$$\tilde{\dot{\rho}}_k \simeq (1 - \frac{\varepsilon}{R})(\rho_0 \gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k) - (1 - \frac{\varepsilon}{R}) \bar{\dot{\rho}}_k - \frac{\varepsilon}{R} v_k \qquad (4.3.12)$$

When b is determined by equation (2.2.14), the steady state gain and variance are also seen not to disappear. Using the covariance recursion, one finds for the steady state,

$$2b - b^2 = 2 + \frac{2\beta R}{P} - 1 - \frac{2\beta R}{P} - \frac{\beta^2 R^2}{P^2}$$

$$= 1 - \frac{\beta^2 R^2}{P^2}$$

$$P = \left[1 - (1 - \beta^2 R^2/P^2) \frac{P}{(P + R)}\right] P$$

$$= \left[P - (1 - \frac{\beta^2 R^2}{P^2}) \frac{P^2}{(P + R)}\right]$$

$$P = \frac{(P + \beta^2 R)}{P + R} R$$

Solving for P, one finds that

$$P = \beta R \qquad (4.3.13)$$

For the steady state gain,

$$K' = bK = (1 + \beta \frac{R}{P}) \frac{P}{P + R}$$

which yields, on substituting for P from (4.3.13),

$$K' = bK = \frac{2\beta}{\beta + 1} \qquad (4.3.14)$$

When $\beta = 0$, one has the optimal Kalman filter, and the steady state values of gain and variance are zero. For $\beta = 1$, $P = R$ and $K' = 1$, as expected.

For very large k, the estimate of $\dot{\rho}$, including the first order model error effects, is found to be approximately

$$\hat{\dot{\rho}}_k \simeq \bar{\dot{\rho}}_k + \frac{2\beta}{\beta + 1} [\rho_0 \gamma \sinh \gamma t_k + \dot{\rho} \cosh \gamma t_k - \bar{\dot{\rho}}_k + v_k]$$

or

$$\hat{\dot{\rho}}_k \simeq \frac{1 - \beta}{1 + \beta} \bar{\dot{\rho}}_k + \frac{2\beta}{\beta + 1} [\rho_0 \gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k + v_k] \qquad (4.3.15)$$

The corresponding error is found to be approximately

$$\tilde{\dot{\rho}}_k = \frac{1-\beta}{1+\beta}\left[\rho_0\gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k\right] - \frac{1-\beta}{1+\beta}\tilde{\dot{\rho}}_k - \frac{2\beta}{\beta+1} v_k \qquad (4.3.16)$$

Again the results are as expected. For $\beta = 0$, the error consists of only the time varying term. For $\beta = 1$, the error is due solely to the measurement noise.

4.4 Schmidt Suboptimal Filter: Additive Gain Term

Using the same procedures as in section 4.3 for $h = 1 + \beta K/P$, identical results are obtained for the steady gain and variance. Using equation *(2.3.6)*,

$$P = \left(1 - \frac{P}{P+R}\right) P + \frac{\beta^2 R^2}{P+R}$$

$$= \frac{(P + \beta^2 R) R}{P+R}$$

or

$$P = \beta R \qquad (4.4.1)$$

For the gain,

$$M = \frac{P}{P+R} + \frac{\beta R}{P+R}$$

or using $P = \beta R$,

$$M = \frac{2\beta}{\beta+1} \qquad (4.4.2)$$

It is noted that the difference between the two algorithms, discussed in section 2.3, does not appear here. In other words, for the gain scaling suboptimal term, the variance term divides out, yielding the identical results in gain and variance.

4.5 Limited Memory Filter

To apply the limited memory filter to the rendezvous problem, assume we have available $\hat{\rho}_{m|m}$ and $P_{m|m}$, and it is desired to obtain the limited memory estimate over the set of $(k - m)$ observations. The necessary quantities required by equations *(2.4.8)* and *(2.4.9)* are given below.

$$\hat{x}_{k|k} = \hat{\rho}_{k|k} = \hat{\rho}_{k|k-1} + K_k \, r_{k|k-1}$$

$$= \hat{\rho}_{m|m} + \sum_{j=m+1}^{k} K_j \, r_{j|j-1}$$

where the predicted residual is

$$r_{j|j-1} = \rho_0 \gamma \sinh \gamma t_j + \dot{\rho}_0 \cosh \gamma t_j - \hat{\rho}_{j|j-1} + v_j \qquad (4.5.1)$$

and the gain at t_j is

$$K_j = \frac{P_{m|m}}{(j-m) P_{m|m} + R} \quad , \quad m + 1 \leq j \leq k \qquad (4.5.2)$$

Similarly,

$$P_{k|k} = \frac{P_{m|m} R}{(k-m) P_{m|m} + R} \qquad (4.5.3)$$

and

$$P_{k|m} = P_{m|m} \qquad (4.5.4)$$

$P_{k|(k-m)}$, the error covariance based on observations from t_m to t_k, may be found directly from equations *(2.4.9)* or *(2.4.10)* or may be computed recursively by *(2.4.12)*. The computation using equation *(2.4.9)* yields

$$P_{k|(k-m)} = [P_{k|k}^{-1} - P_{k|m}^{-1}]^{-1}$$

$$= \left[\frac{(k-m) P_{m|m} + R}{P_{m|m} R} - \frac{1}{P_{m|m}} \right]^{-1}$$

or

$$P_{k|(k-m)} = R/(k-m) \tag{4.5.5}$$

It is noted in passing that for $k-m = 1$, i.e., one observation, the corresponding error covariance is as expected, $P_{k|k-m} = R$.

Using equation *(2.4.8)* the limited memory estimate is found.

$$\hat{\dot{\rho}}_{k|(k-m)} = \frac{R}{(k-m)} \left\{ \frac{(k-m)\,P_{m|m} + R}{P_{m|m}\,R} \left[\hat{\dot{\rho}}_{m|m} + \sum_{j=m+1}^{k} \frac{P_{m|m}\,r_{j|j-1}}{(j-m)\,P_{m|m} + R} \right] - \frac{\hat{\dot{\rho}}_{m|m}}{P_{m|m}} \right\}$$

$$= \frac{R}{(k-m)} \left\{ \frac{(k-m)}{R}\,\hat{\dot{\rho}}_{m|m} + \frac{(k-m)P_{m|m} + R}{R} \sum_{j=m+1}^{k} \frac{r_{j|j-1}}{(j-m)\,P_{m|m} + R} \right\}$$

or

$$\hat{\dot{\rho}}_{k|(k-m)} = \hat{\dot{\rho}}_{m|m} + \left[P_{m|m} + \frac{R}{(k-m)}\right] \sum_{j=m+1}^{k} \frac{r_{j|j-1}}{(j-m)\,P_{m|m} + R} \tag{4.5.6}$$

Thus, the outputs of *(4.5.5)* and *(4.5.6)* are the limited memory error covariance and estimate. Note that when $k-m = 1$, the estimate depends only upon the most recent observation. In this case *(4.5.6)* gives

$$\hat{\dot{\rho}}_{k|(1)} = \hat{\dot{\rho}}_{k-1|k-1} + r_{k|k-1}$$

$$= \rho_0 \gamma \sinh \gamma t_k + \dot{\rho}_0 \cosh \gamma t_k + v_k \tag{4.5.7}$$

4.6 Numerical Results

The foregoing sections have considered the application of some of the error compensation methods to the rendezvous problem. In particular, various closed form expressions for the range-rate estimate, error, and steady-state gain and covariance equations were derived. Such expressions help provide analytical insight into the filter structures. However, the performance of the algorithms can be seen best, perhaps, through numerical simulations.

Thus, in this section numerical results are obtained for each of the compensation algorithms previously presented.

To carry out the simulations, nominal parameter values defining the dynamic model were selected as shown in Table 4.1. The model error parameter, γ, is defined by equation *(1.4.20)* as

$$\gamma^2 = \frac{3\mu}{|r_t|^3} \sin^2 \varepsilon$$

and reflects the gravitational parameter, μ, the target vehicle radial distance, r_t, and the line-of-sight angle (above the local horizontal). For Earth orbit, $\mu = 398603.2\ km^3/sec^2$. Thus, using the value of γ given in the table, values of ε are determined for various values of $|r_t|$. For example, an orbital altitude of 160 km ($|r_t| = 6538.165$ km) corresponds to $\varepsilon = 4.96°$; at 240 km, $\varepsilon = 5.1°$.

Parameter	*Value*
Initial range, ρ_0	*10 km*
Initial range rate, $\dot{\rho}_0$	*-1 m/sec*
Initial range rate error, $\tilde{\dot{\rho}}_0$	*0 m*
A priori variance, P_0	*10* $(m/sec)^2$
Observation error variance, R	*.01* $(m/sec)^2$
Observation interval, Δt	*10 sec*
Model error parameter, γ	$.18 \times 10^{-3}\ (sec)^{-1}$

Table 4.1 Nominal Simulation Parameters

In performing the simulations, the nominal or assumed model, as before, is $\dot{\rho}_k = \dot{\rho}_{k-1}$. The true dynamics are computed according to the linearized model, equation *(1.4.22)*. The true range rate observations are computed by adding a random number, distributed* $N(0, R)$, to the true range rate from *(1.4.22)*. The observation deviation is then found by *(1.4.23)*, and the filtering equations of the various algorithms are applied to obtain the gain, estimate and updated state error variance. In all cases investigated, the true initial error is zero, as seen from Table 4.1. Also, to provide a common basis for evaluation, the error sequence is identical in all cases.

Standard Minimum Variance Estimates. To provide some standards of reference, Figure 4.1 shows the behavior of the ordinary minimum variance algorithm in the presence of model error with no model error compensation techniques employed. Also shown is the case with perfect modelling. The ordinate axis represents the estimate error in meters, with the abscissa measuring time in seconds. Three curves are shown. The solid line represents the actual estimate error, $\tilde{\dot{\rho}}_k$, where the model error *(1.4.25)* is present. The dotted line gives the estimate error for no modelling error. The dashed line represents the computed error standard deviation (s.d.), $\sigma_k \equiv \sqrt{P_k}$. Note that this curve is the same for both cases, regardless of the presence or absence of model errors, since the estimation algorithms are identical.

For the incorrect model, the rapid rate at which the estimate of the range rate diverges is rather astonishing. Thus at typical Earth orbit

---

* The notation $N(0, R)$ means that a random number is sampled from a normal distribution of zero mean and variance, R. In the simulations, normally distributed random numbers were calculated using the Central Limit Theorem (46, 98-99). Further, the same sequence of random numbers was repeated from run to run.

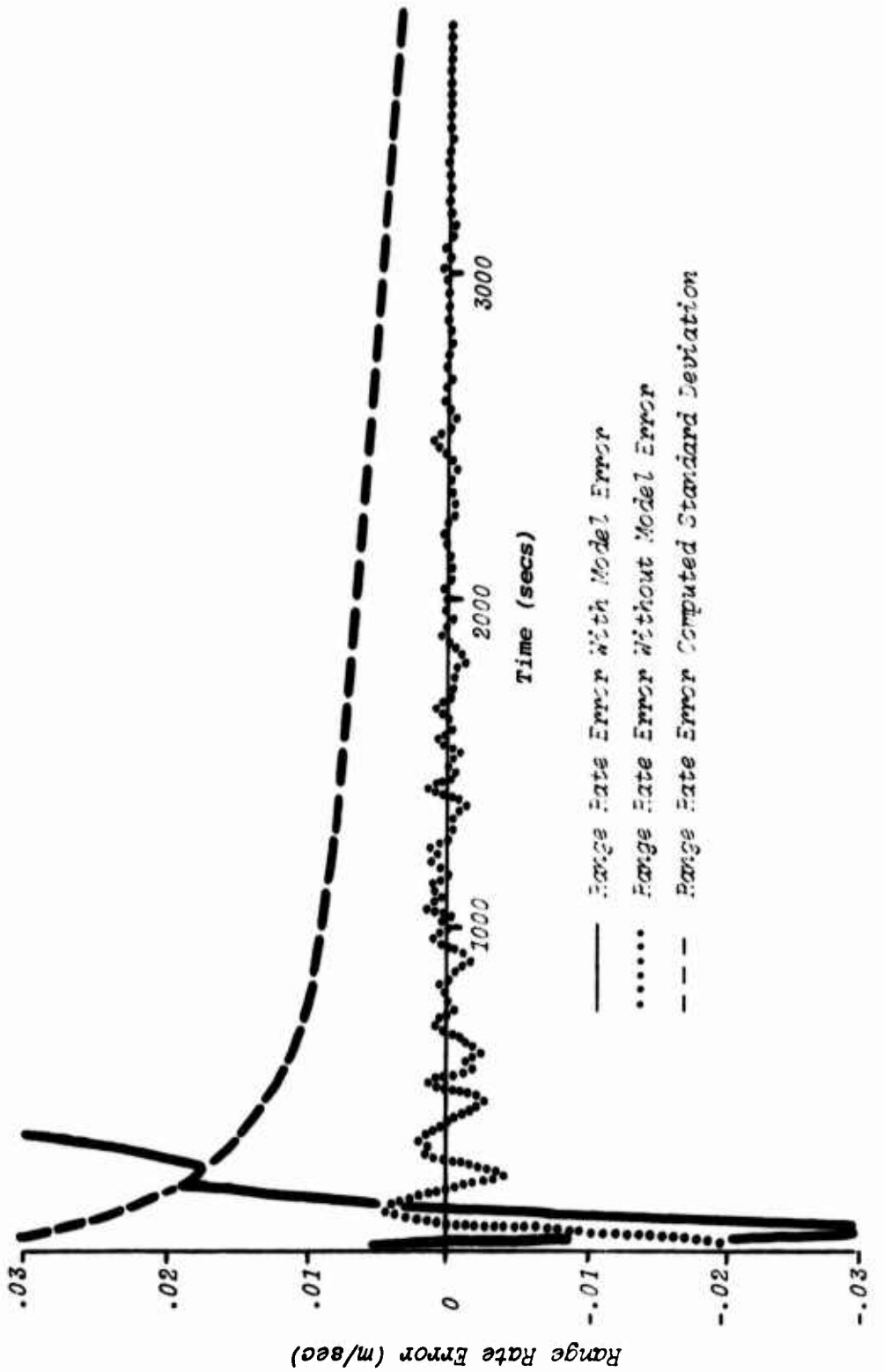


Figure 4.1 Range Rate Error With and Without Model Error

altitudes, relatively small line-of-sight angles -- on the order of 5° -- produce a dramatic divergence of the range rate estimate. On the other hand, for the exact model *(1.4.22)* the power of the Kalman filter is reflected by the dotted curve: after 1500 seconds the error has practically disappeared. Thus these two curves can be seen to bound the performance of any of the compensation algorithms investigated for this problem.

Commonality of the Algorithms. Because of the fact we are dealing with a single state variable, it is possible to relate most of the suboptimal filter parameters to obtain equivalent steady stage filter performance. In other words, through appropriate expressions relating the filter parameters to each other, it is possible to obtain a value for each parameter which will produce equivalent steady state performance for each algorithm. (Exceptions, which are treated separately, are the ordinary [b = constant] gain scaling algorithm and the limited memory filter.)

Consider the steady state relations for the age-weighting, modified gain scaling, and additive gain term algorithms. These are given by *(4.2.5)*, *(4.3.13)*, and *(4.3.17)*, or *(4.4.1)* respectively as

$$P_\infty = \frac{s - 1}{s} R$$

$$P_\infty = \frac{\alpha R}{? - \alpha}$$

$$P_\infty = \beta R$$

(Recall that this last equation is the same for the modified gain scaling of equation *(2.2.14)* and for the additive gain term.) For equivalent steady state performance each of the above expressions must be equal to any other, thus providing the necessary relationships relating the filter parameters. Therefore,

$$\beta = \frac{\alpha}{2 - \alpha} = \frac{s - 1}{s} \qquad (4.6.1)$$

For the case where the state noise covariance is employed, the steady state expression given by *(1.4.32)*, may be equated to the corresponding expressions containing the suboptimal filter parameters. Thus,

$$P_\infty = \frac{1 + \sqrt{1 + 4/f}}{1 + \sqrt{1 + 4/f} + 2/f} R \qquad (1.4.32)$$

where $f = Q/R$, may be included in *(4.6.1)* so that

$$\beta = \frac{\alpha}{2 - \alpha} = \frac{s - 1}{s} = \frac{1 + \sqrt{1 + 4/f}}{1 + \sqrt{1 + 4/f} + 2/f} \qquad (4.6.2)$$

Using these equations, a table of values may be constructed. Table 4.2 gives such appropriate values which yield the equivalent filter performance. The table was constructed by selecting various values of β and then solving for the corresponding values of α, s, and Q.

β	α	s	Q	P_∞
0.	0.	1.0	0.	0.
.10	.1818	1.111	.0111R	.1R
.20	.3333	1.250	.0500R	.2R
.30	.4615	1.429	.1286R	.3R
.40	.5714	1.667	.2667R	.4R
.50	.6667	2.000	.5000R	.5R
.60	.7500	2.500	.9000R	.6R
.70	.8235	3.333	1.633R	.7R
.80	.8889	5.000	3.200R	.8R
.90	.9474	10.00	8.100R	.9R
.99	.9950	100.0	98.01R	.99R
1.00	1.000	∞	∞	R

Table 4.2 Filter Parameter Values for Equivalent Steady State Performance

Figures 4.2a through 4.2e illustrate the numerical equivalence in terms of filter performance. The curves were actually generated using the additive gain term algorithm for $\beta = .1, .3, .6, .9$, and 1.0, respectively. However, steady state conditions, and hence equivalent performance, are quickly reached. These curves therefore illustrate essentially the same performance for each of the algorithms. As in Figure 4.1, the solid line represents the actual estimate error, $\tilde{\rho}_k$. The dashed lines represent plus and minus values of the computed standard deviation, σ_k.

In examining the figures one finds that the maximum and minimum error values increase with increasing values of the filter parameters. In the extreme (Figure 4.2e), the estimate follows the observations, and the errors are greatest in absolute value. Based on the sample of 384 points, 71.61% of the estimates are within the 1σ value* of .1 meter, determined from the error covariance computed by the filter. In Figure 4.2a the absolute value of the maximum and minimum errors are smallest. Here, 64.06% of the 384 points are within the 1σ value of .03163. However, the model error appears to be showing itself in the form of a bias; the majority of the error values lie above the expected value of zero. In Figure 4.2b the error does not exhibit such a bias, thus suggesting parameter values somewhere between those for Figures 4.2a and 4.2b (e.g., $.1 \leq \beta \leq .3$) -- at least for the time inverval of 3840 seconds. However, regardless of the value selected, as long as $\beta < 1$, divergence can be expected to occur eventually. This follows from the unboundedness of the error, equation *(4.3.20)*.

Because the steady state variance associated with the ordinary gain scaling (b = constant) is zero, the equivalence developed above cannot be

---

* A short discussion of an interpretation of σ in light of normal ergodic processes is given in Appendix D.

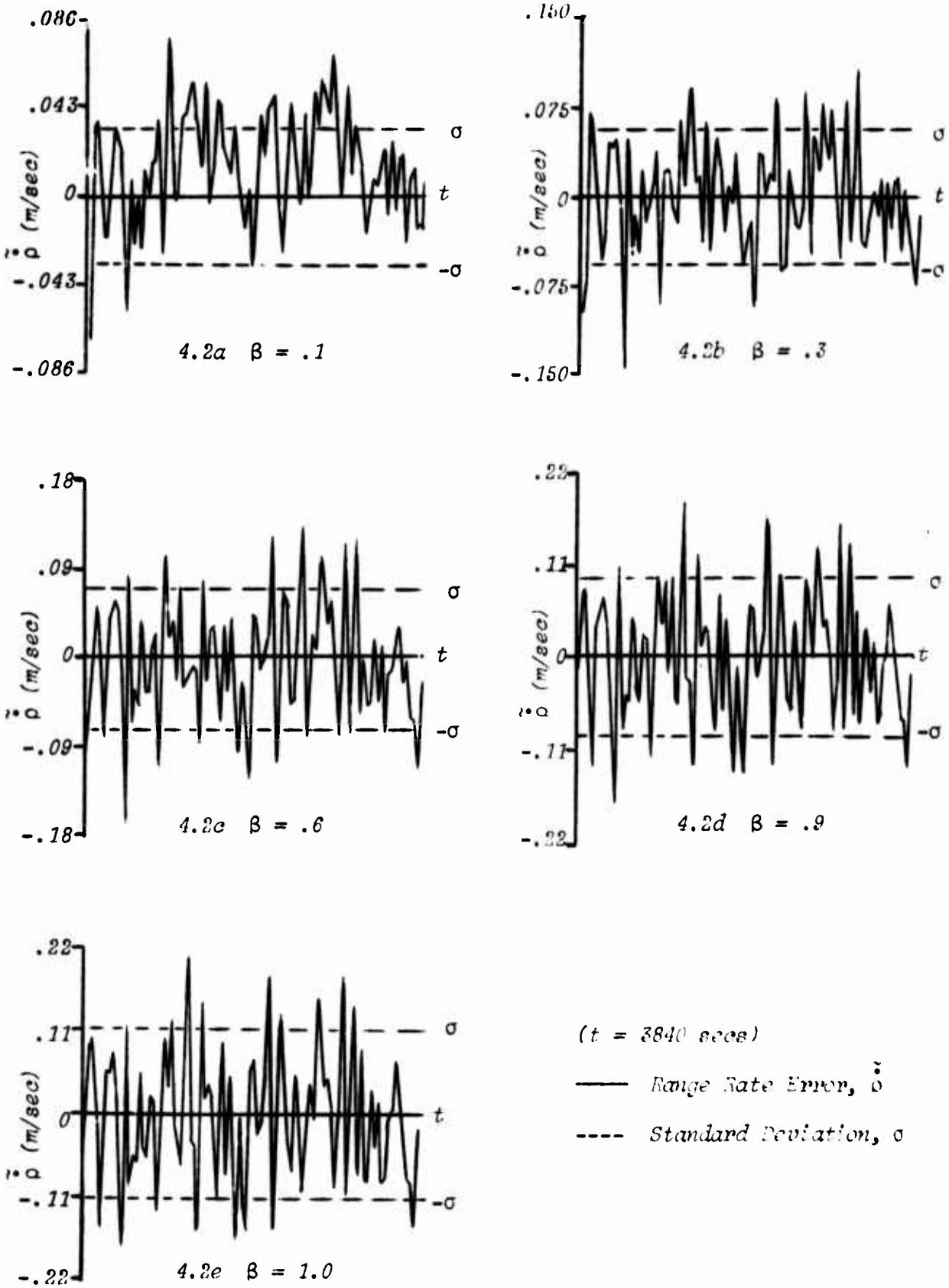


Figure 4.2 Equivalent Filter Performances for β, α, ε, and Q

obtained. Thus this algorithm results in a completely different, and not very effective, filter performance. This is seen in Figure 4.3, which shows the estimate errors for the gain scaling algorithm using values of b = 1.3, 1.6, and 1.9, respectively. Clearly the performance is unacceptable; divergence occurs soon in all cases. The values b = 1.3 and b = 1.6 are seen to have little effect on the estimation performance. Only when b becomes close to 2 does the algorithm have an effect in delaying the divergence. Even then, the variance eventually becomes zero, although it simply takes longer. Recall from equation *(4.3.12)* that b = 2 prevents any change in the variance, and hence $P_\infty = P_0$. For this problem the ordinary gain scaling is obviously not acceptable.

Adaptive Estimation of State Noise Covariance and Suboptimal Parameters. Numerical simulations were performed for each of the adaptive forms for estimating the state noise covariance and suboptimal parameters, as presented in sections 3.1 and 3.2. In these simulations, the values of γ given by *(3.1.13)* and *(3.1.14)* were used (the sample mean of the square of the previous N predicted residuals. Initially, *(3.1.14)* was used to "build up" the set of N residuals. Subsequently *(3.1.13)* was used to compute γ_N. Use of equation *(3.1.12)* was not investigated due to the inherent disadvantage of either having the estimate lag the observations, or having to reprocess or smooth the estimates.

In computing the various parameters, the following equations were employed:

$$Q_k = max\ [0.,\ \gamma_N - P_K - R] \qquad (4.6.3a)$$

$$s_k = max\ [1.,\ (\gamma_N - R)/P_k] \qquad (4.6.3b)$$

$$b_k = max\ [1.,\ (\gamma_N - R)(P_k + R)/P_k] \qquad (4.6.3c)$$

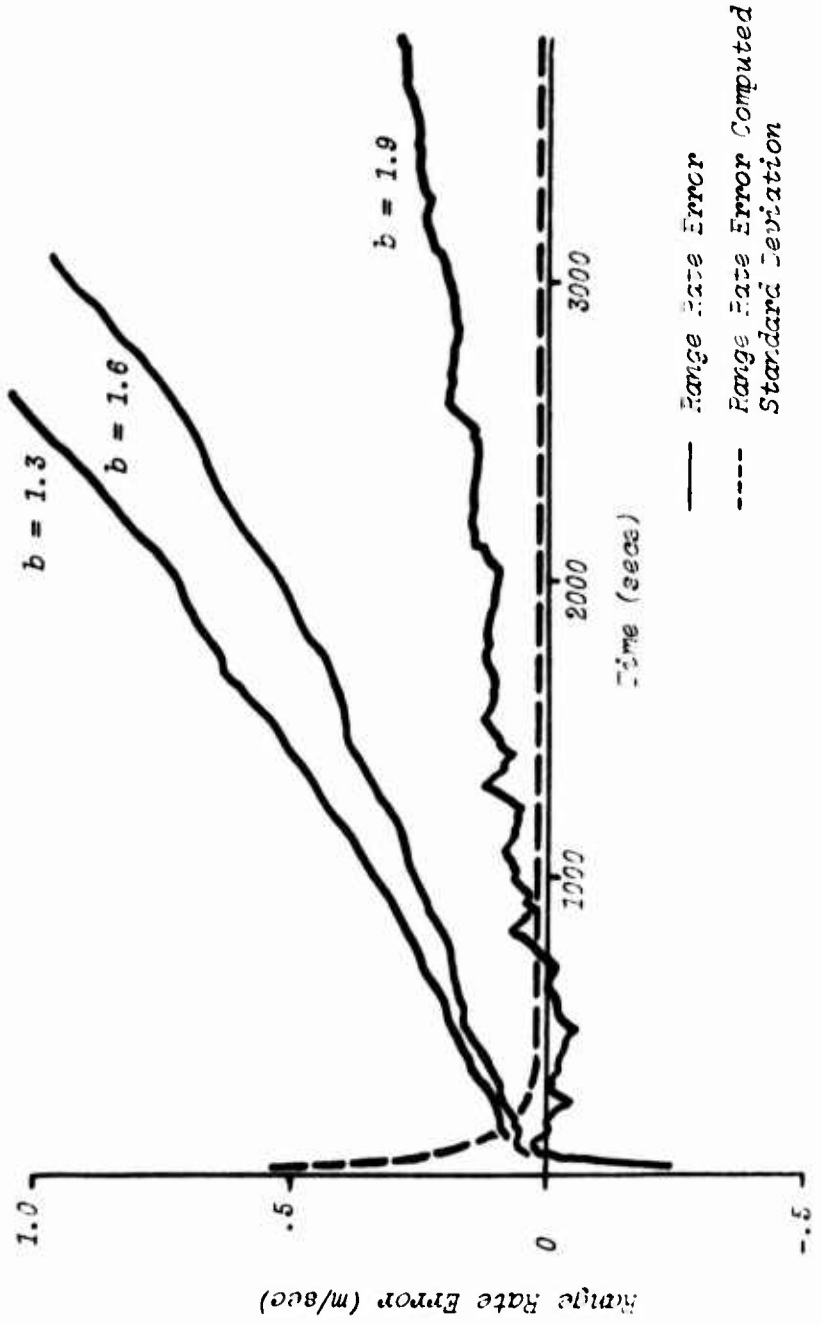


Figure 4.3 Range Rate Error for Standard Gain Scaling

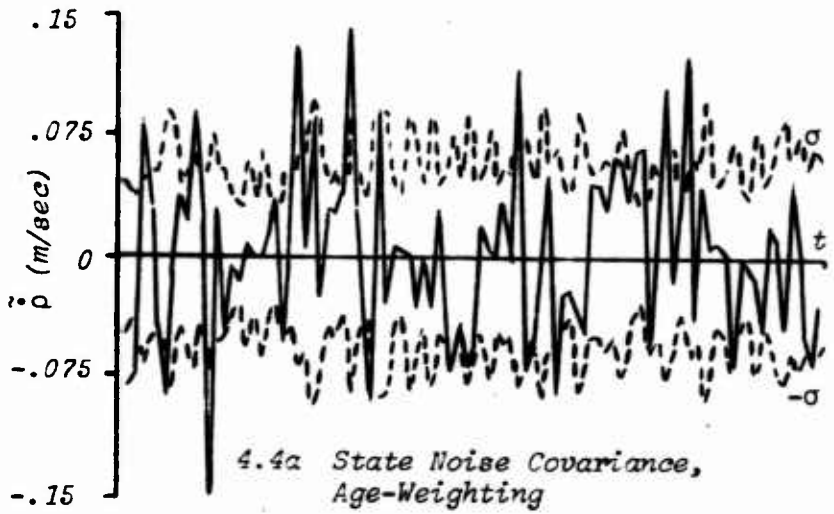
$$\alpha_k = max\ [(\gamma_N - R)/\gamma_N,\ P_k/(P_k + R)] \qquad (4.6.3d)$$

$$\beta_k = 1. - (P_k + R)/max\ [(P_k + R),\ \gamma_N] \qquad (4.6.3e)$$

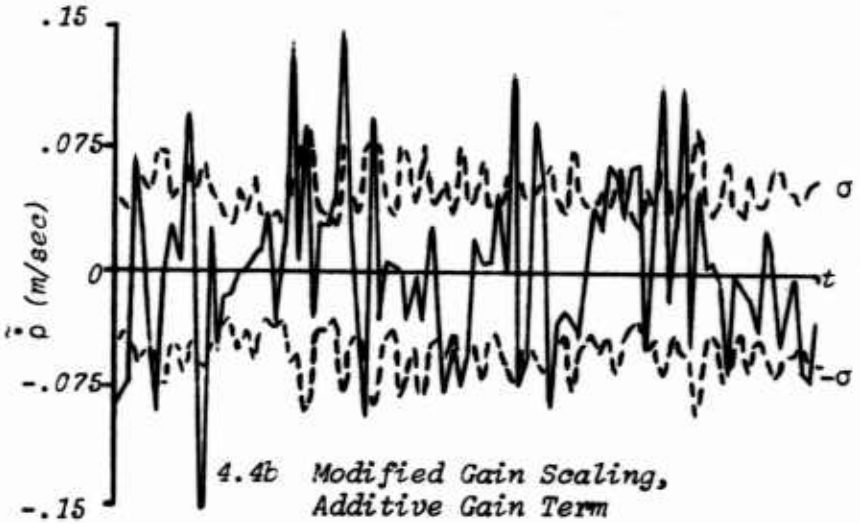
These equations yield appropriate values for the parameters according to the values of γ_N. They also automatically incorporate the attendant limits on the parameter values, which is the reason for the presence of the *max* function.

In carrying out the simulation, some particularly interesting, if not surprising, results were obtained. Three different values of N were used in computing γ_N, namely $N = 1$, 10, and 20. In each case, the algorithms exhibited very similar behavior, with the exception of the ordinary gain scaling algorithm. In fact, it was found that the performances of Jazwinski's adaptive Q-estimator yielded results identical with the adaptive age-weighting algorithm. Further, identical results were also obtained with the adaptive forms of the two modified gain scaling algorithms and the additive gain term approach of Schmidt. However, some small differences were noted between these two sets of performances. The simulation results are shown for each N in Figures 4.4, 4.5, and 4.6, respectively.

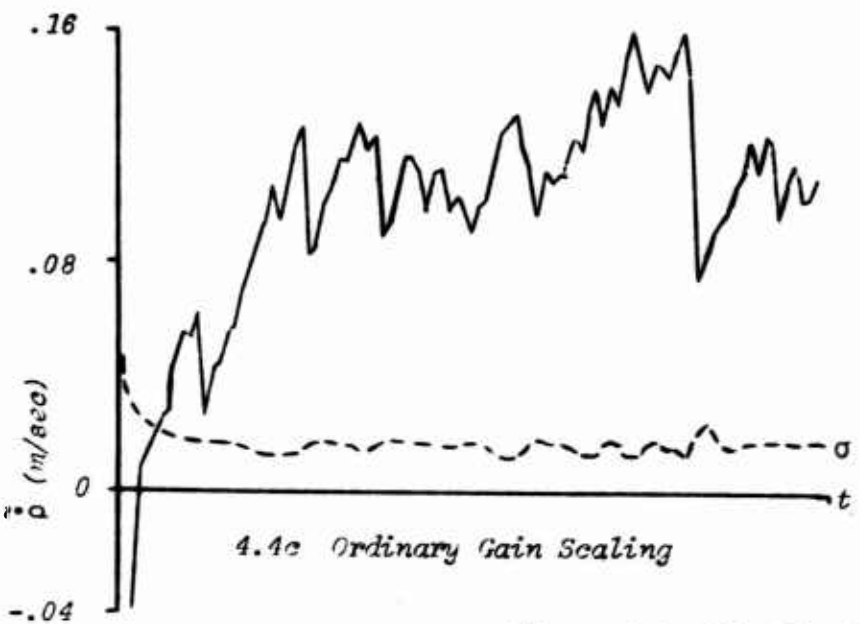
Figure 4.4a shows the estimate error and $\pm 1\sigma$ curves for Jazwinski's Q-estimator and adaptive age-weighting with $N = 1$. Here, 64.84% of the 384 estimates are within the computed 1σ of the true value. Very similar estimate error behavior is shown in Figure 4.4b for the modified gain scaling algorithms and additive gain term technique. Although the performances are nearly the same for the two sets of plots, in this latter case only 59.35% of the 384 estimates are within 1σ of the true value. On this basis the adaptive forms of the state noise covariance and age-weighting algorithms appear to offer greater confidence in their associated estimates. This suggestion seems to be further substantiated by Figures 4.5a and 4.5b. Here the same type of information is plotted as in Figures 4.4, but now $N = 10$. The



4.4a State Noise Covariance, Age-Weighting



4.4b Modified Gain Scaling, Additive Gain Term



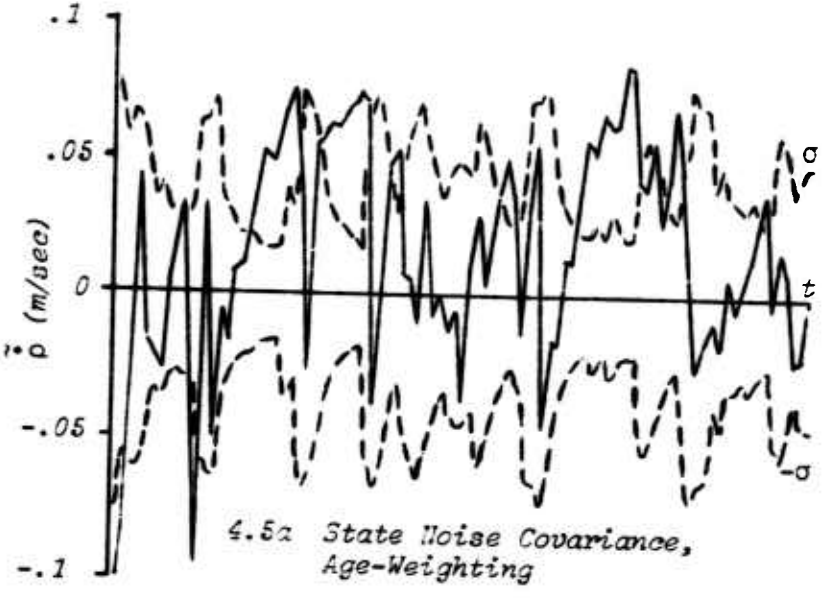
4.4c Ordinary Gain Scaling

—— Range Rate Error, $\tilde{\dot{\rho}}$

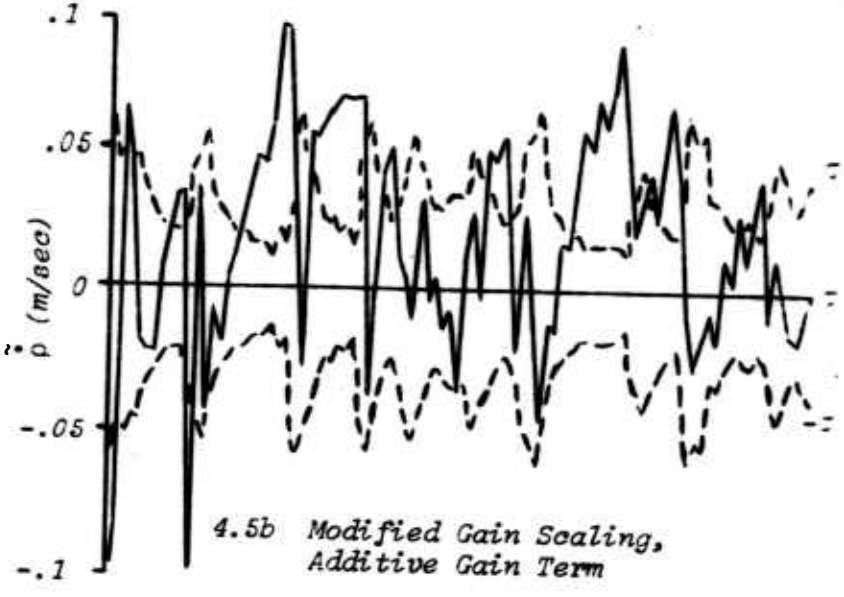
---- Standard Deviation, σ

t = 3840 secs

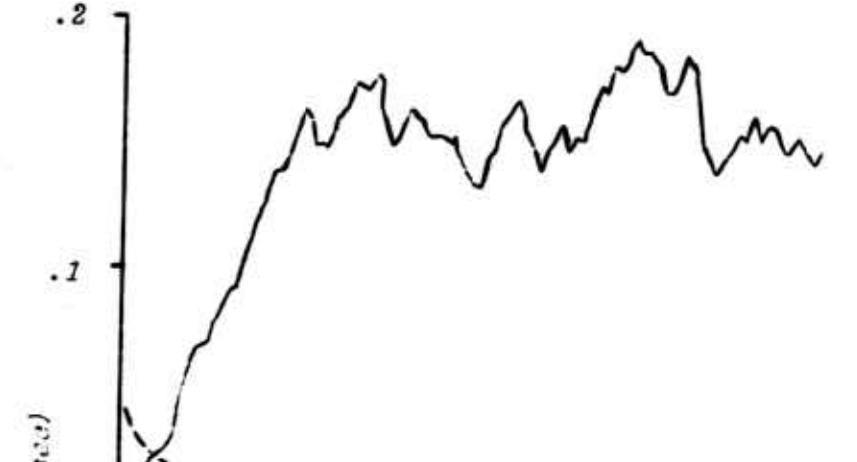
Figure 4.4 Adaptive Algorithms for 1 Residual



4.5a State Noise Covariance, Age-Weighting



4.5b Modified Gain Scaling, Additive Gain Term



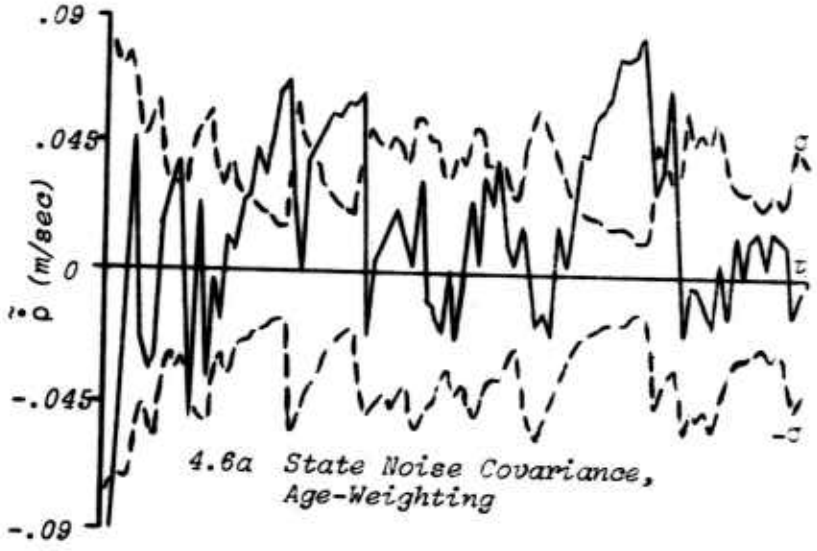
4.5c Ordinary Gain Scaling

——— Range Rate Error, $\dot{\tilde{\rho}}$

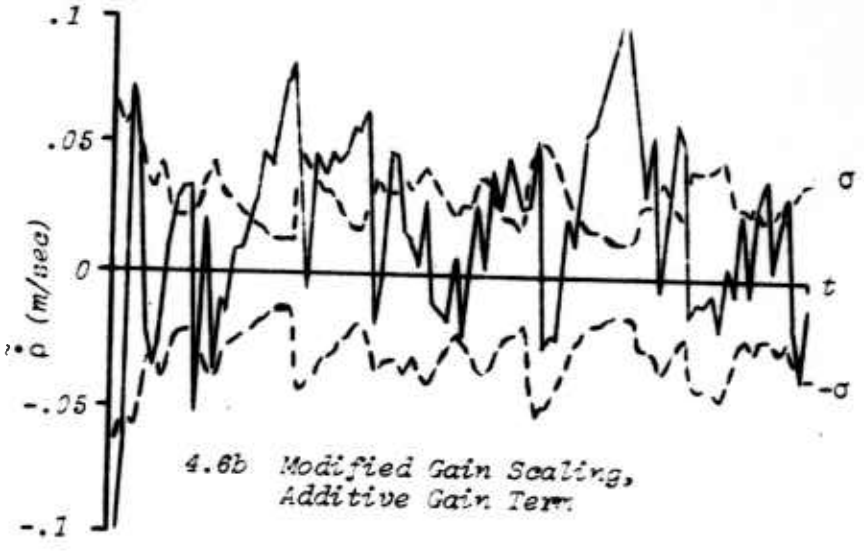
------ Standard Deviation, σ

t = 3840 secs

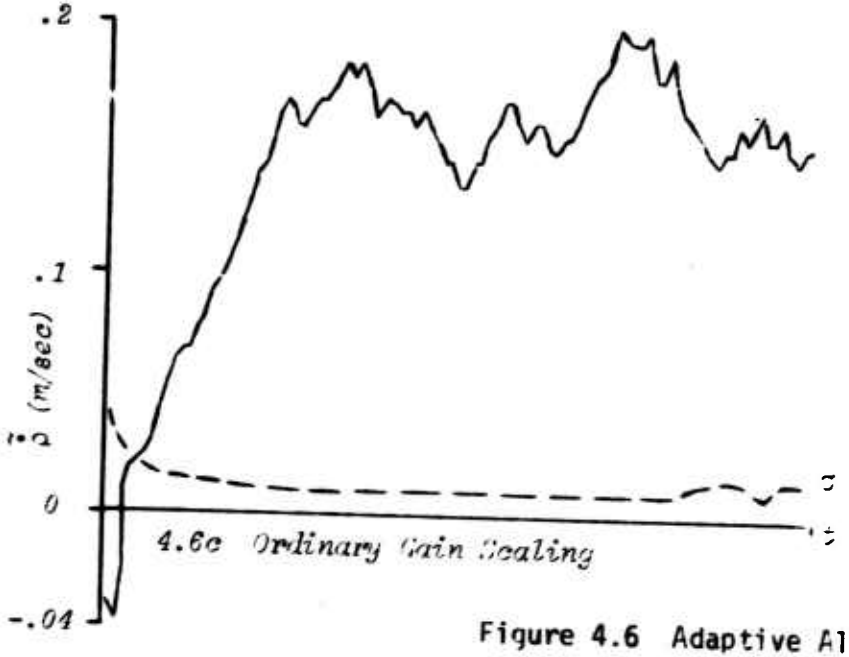
Figure 4.5 Adaptive Algorithms for 10 Residuals



4.6a State Noise Covariance, Age-Weighting



4.6b Modified Gain Scaling, Additive Gain Term



4.6c Ordinary Gain Scaling

—— Range Rate Error, $\dot{\tilde{\rho}}$

---- Standard Deviation, σ

t = 3840 secs

Figure 4.6 Adaptive Algorithms for 20 Residuals

percentages of the 384 points lying within $\pm 1\sigma$ of the mean are 55.99% and 50.78% respectively for Figures 4.5a and 4.5b. The same trend is present for the case where $N = 20$, Figures 4.6a and b. Here the respective percentages are 59.9% and 51.30%. Thus the implication is that the former algorithms offer greater confidence, as measured by the computed covariance.

Another notable characteristic is present in each of Figures 4.5a and 4.5b. Close examination suggests the appearance of the model error bias, since the majority of the error values appear above the expected zero value on both sets of plots. This implies that when more residuals are used in forming γ_N, the bias errors have more opportunity to affect the estimate. Increasing the residual sample size to $N = 20$ yields the plots shown in Figures 4.6 which confirm this suspicion. In retrospect this should not be surprising; as $N \to \infty$ the adaptive algorithms cease to operate, and the bias errors dominate producing divergence of the estimate.

For each of the values of $N = 1, 10, 20$, the adaptive form of the ordinary gain scaling algorithm yields the results plotted in Figures 4.4c, 4.5c, and 4.6c. Clearly these results are unacceptable. Initially the error variance is large; hence the adaptive algorithm obtains no information from the residuals since γ_N is generally small and hence *(4.6.3c)* produces $b_k = 1$. As the process continues, γ_N tends to increase, but the sensitivity of the algorithm is simply not sufficient to control the divergence of the estimate. There appears to be a simple explanation for this. For the adaptive form of each algorithm consider the sensitivity of the gain to changes in the residual sample, γ_N. This is determined by using the appropriate member of equations *(4.6.3)* in forming the corresponding gain, and then taking the partial derivative with respect to γ_N. In all but the case of the ordinary gain scaling algorithm, the partial derivative is

$$\frac{\partial K}{\partial \gamma_N} = \frac{K}{\gamma_N^2} \qquad (4.6.4)$$

For the ordinary gain scaling,

$$\frac{\partial K}{\partial \gamma_N} = \frac{1}{P} \qquad (4.6.5)$$

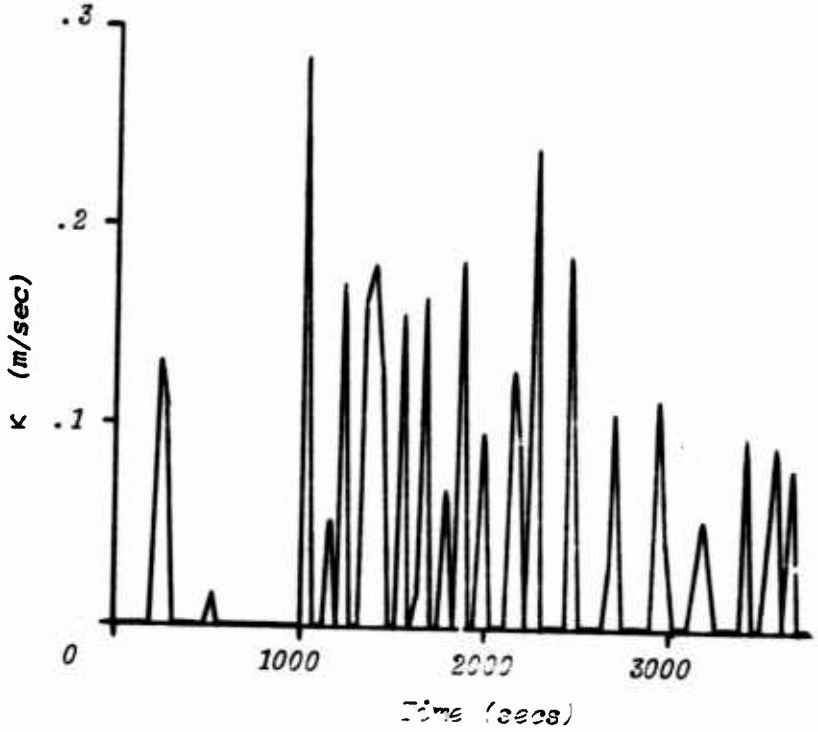
From these equations the reason for the very different behavior should be apparent. First, note that when the residuals are very small the sensitivity given by *(4.6.4)* is very high. This is desirable since the gain values then respond quickly to changes in γ_N. However, for gain scaling, a similar sensitivity requires that P be very small. This is exactly the type of behavior displayed in Figure 4.6c. Only when the error variance has decreased significantly does the adaptive feature begin to operate. Thus based upon the indicated performance, the ordinary gain scaling algorithm will be dismissed from further consideration as a viable filtering algorithm.

No discussion has been given yet concerning the computed values of the state noise variance or suboptimal filter parameters. An obvious approach is simply to "shotgun" the reader with a set of some 21 graphs of estimation curves and then pick through the data a "pellet" at a time. A preferred approach, and the one followed here, is to offer a representative sample and then to note certain trends determined from examination of this and previous data. In this manner we hope to avoid "blowing the reader's mind" * and still provide a further measure of insight into the behavior and relationships of the algorithms.

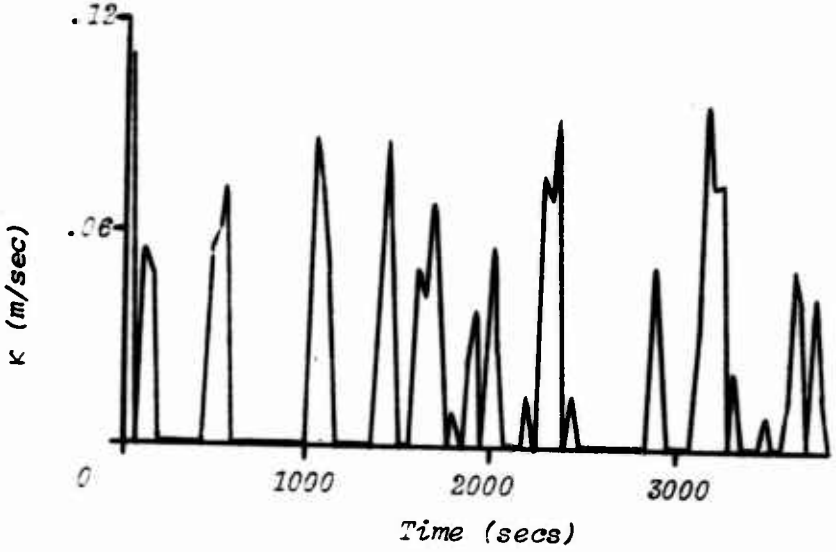
We proceed by offering, as the representative sample, plots of the state noise standard deviation, $\kappa \equiv \sqrt{Q}$, given in Figures 4.7a through c. Figure 4.7a gives the values of κ determined from Jazwinski's Q-estimator

---

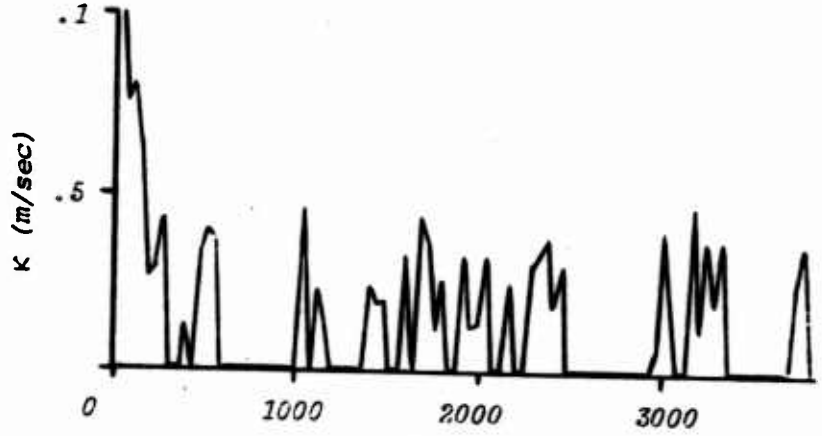
* This obvious pun could not be resisted.



4.7a 1 Residual



4.7b 10 Residuals



4.7c 20 Residuals

Figure 4.7 State Noise Standard Deviation, κ, for Various Residual Sample Sizes

based upon single residuals ($N = 1$ for γ_N). During the initial phase of operation, while the state error covariance is reasonably large, few non-zero values of Q are computed. After 800 seconds, the error covariance has decreased, and there is a high frequency of non-zero estimates. In Figure 4.7b, the values of κ are shown based upon a 10-residual sample size ($N = 10$ for γ_N). Here both the frequency and magnitude of computed values of κ are less than for the single residual case. Figure 4.7c continues the trend. In all cases the estimated values of Q clearly result in the desired increases of the 1σ curves of Figures 4.4a, 4.5a, and 4.6a. Note, however, that as the residual sample size, N, increases, the state estimate bias error becomes more prominent. As pointed out earlier, as $N \to \infty$, we eventually have no Q-estimates being produced, and thus the state estimate diverges.

For the suboptimal parameters, excluding ordinary gain scaling as stated earlier, similar results have been consistently obtained. Of course, the limiting values are different; however, the parameter estimates show the same random "spike" behavior as those for κ. Similarly, the frequency and magnitude of the values decrease with increasing N, resulting in eventual dominance of the bias error in the state estimate. Further evidence of similar performance for the suboptimal parameter adaptive estimation is given by the similarities of the error curves in Figures 4.4-4.6. If significant differences existed in the adaptive computation, these would also be manifested as significant differences in the various error and 1σ curves. On this basis one may conclude -- and the estimated parameter data bear this out -- that no fundamental differences exist in the performance of Jazwinski's adaptive state noise covariance estimator and the remaining adaptive suboptimal parameter algorithms for this scalar example. (This last qualification is important and will be discussed in section 4.7.)

Limited Memory Filter. Figures 4.8a through 4.8c show the estimate error for Jazwinski's limited memory filter. The figures correspond to the values of the "memory length," i.e., the number of observations processed between applications of the limited memory updates. Thus memory lengths of 1, 10, and 20 observations are reflected in the performance given by Figures 4.8a, b, and c, respectively.

In Figure 4.8a the filter follows the observations. Comparison with Figure 4.2e clearly indicates identical behavior, as expected, with the corresponding steady state covariance, $P_\infty = R$.

For $N = 10$ the estimate error is plotted in Figure 4.8b. In this case the bias error appears as an excess of points above the time axis. However, toward the end of the interval, the central tendency is back toward zero mean. To investigate this further, the case of a 20-observation memory was run. The results appear in Figure 4.8c. As expected, the bias effects are even more pronounced. Indeed, if the memory size is made as long as the complete observation period, the performance is simply that obtained for no error compensation; i.e., the estimate error grows without bound.

In both 4.8b and c the typical behavior of the covariance may be noted. This is expressed as the saw-tooth form of the $\pm 1\sigma$ curves. Each peak corresponds to an increase in the variance produced by the limited memory updated. After each update, the variance begins decreasing in accordance with the operation of the ordinary Kalman filter equations. The width of each of the saw-tooth pulses is therefore directly proportional to the memory length.

The accuracies obtained with the limited memory filter are generally on the order of those obtained by the adaptive forms of the suboptimal

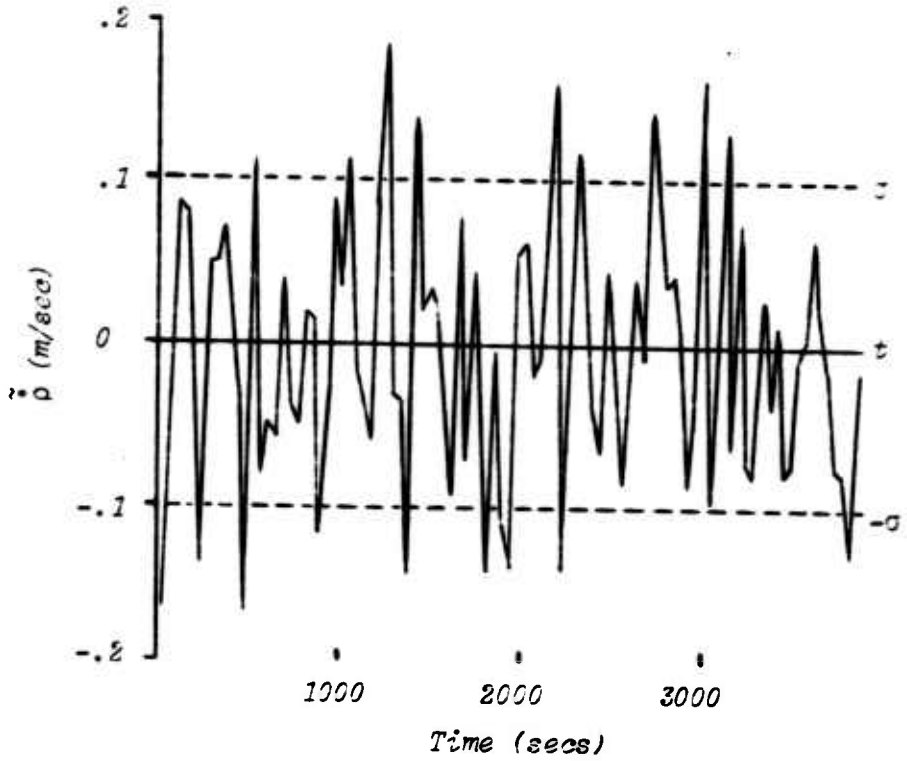
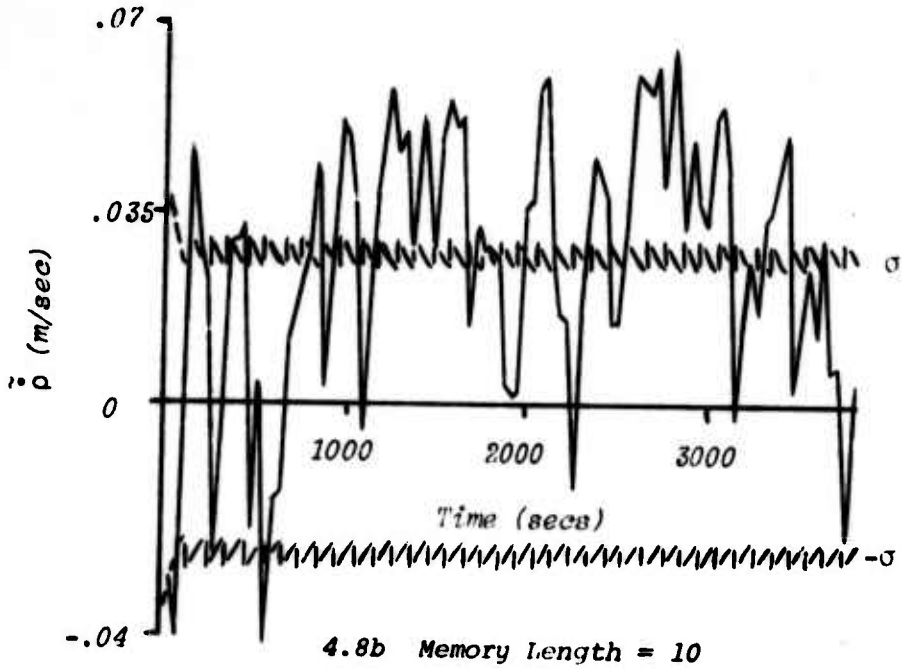
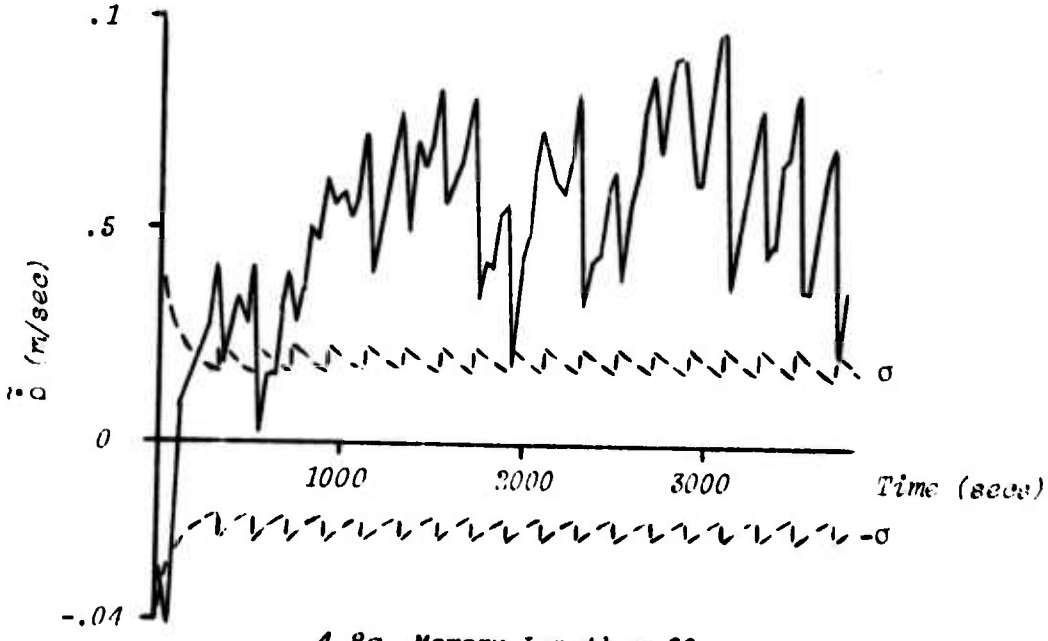


Figure 4.8a Range Rate Error, $\tilde{\dot{\rho}}$, for Limited Memory Filter, Memory Length = 1



4.8b Memory Length = 10



4.8c Memory Length = 20

Figure 4.8b,c Range Rate Error, $\dot{\tilde{\rho}}$, for Limited Memory Filter

algorithms. However, based on the numerical results, the memory length must be smaller than the residual sample size used in the adaptive algorithms.

For this scalar problem the limited memory filter requires no more computational effort than the other adaptive forms. In vector problems involving many state elements, the computational load becomes quite severe due to the required matrix inverses. For this reason the limited memory filter loses its attractiveness for use in more complex problems.

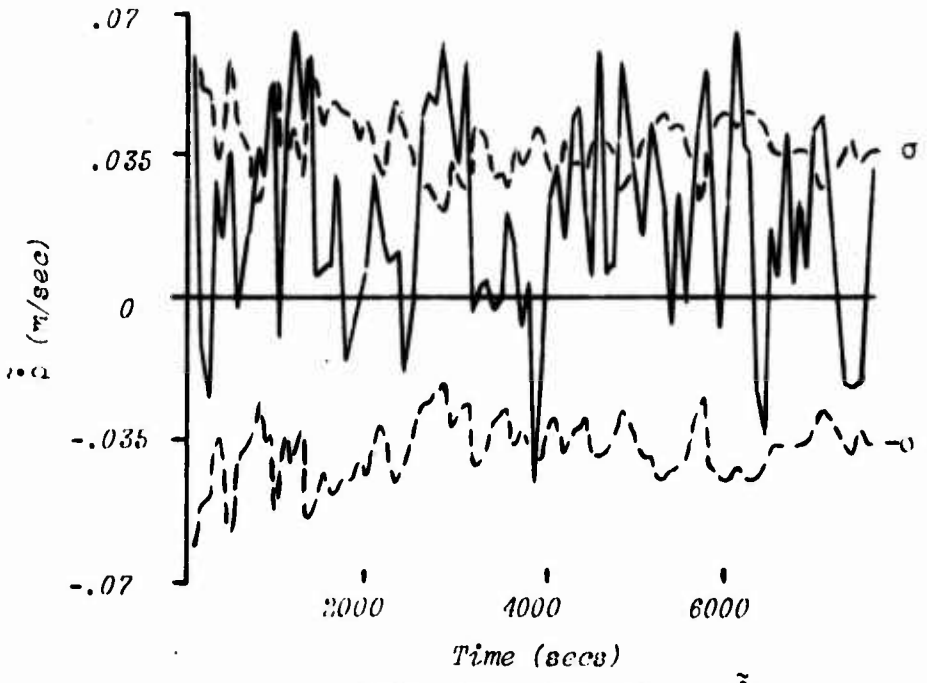
Sequential Estimation of the State Noise Covariance. We turn now to the performance of the sequential estimator for Q. Initial runs of 3840 seconds were made using *a priori* values for the error variance, $S = E\{\tilde{\xi}^2\}$, the residual error variance, $T = E\{\theta^2\}$, and the initial value of Q as follows:

$$S_0 = 1.0 \ (m/sec)^4$$

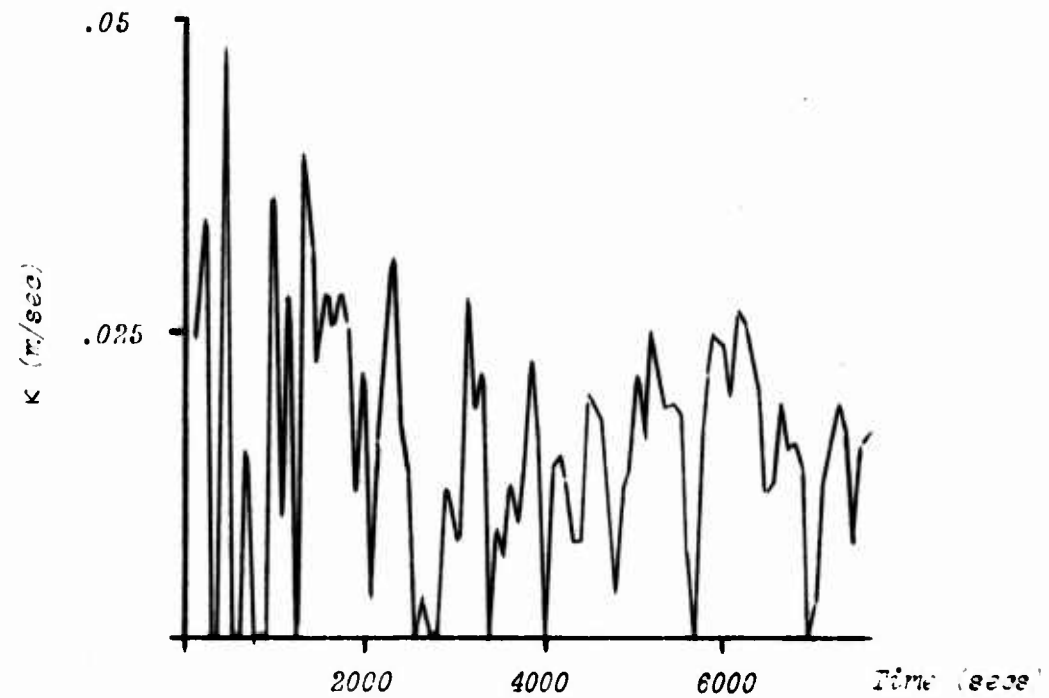
$$T_0 = \bar{\Delta}_1^2 = (y_1 - \bar{\dot{\rho}}_1)^2$$

$$Q_0 = 0. \ (m/sec)^2$$

The performance was slightly better than for Jazwinski's adaptive Q-estimator, although the bias error was apparent. The estimated values of Q took some time to "settle." This appeared to occur near the end of the interval. To investigate this further, the interval was doubled to 7680 seconds; the results are shown in Figures 4.9a and b, which show the estimate errors for $\dot{\rho}$ and estimates of Q respectively. Also shown are the corresponding $\dot{\rho}$-error 1σ curves. The 1σ value for the $\dot{\rho}$ error appears to have a mean of approximately .04 m/sec, with 70.05% of the error being within 1σ. The corresponding value of $\kappa = \sqrt{Q}$ is around .015 m/sec. At the final time, the confidence in the estimate of Q is reflected by a 1σ value of $\sqrt{S} = .027$ $(m/sec)^2$. The data indicated that all the Q-estimates lie under the $\sqrt{S}$ curve, suggesting a decrease in the *a priori* S. Though not immediately obvious, close examination



4.9a Range Rate Error, $\tilde{\dot{\rho}}$

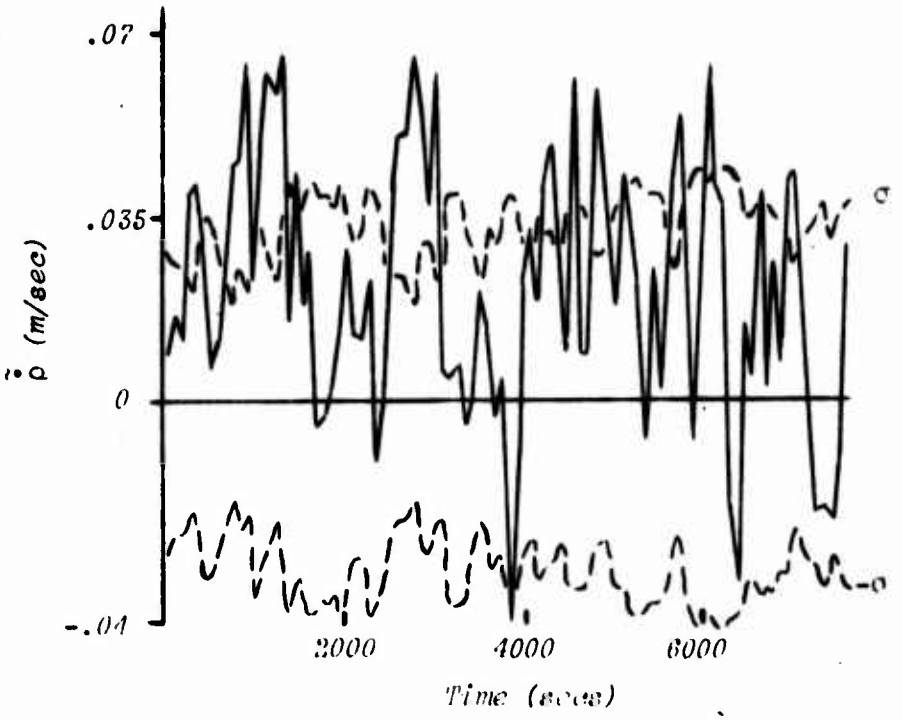


4.9b Standard Deviation, κ

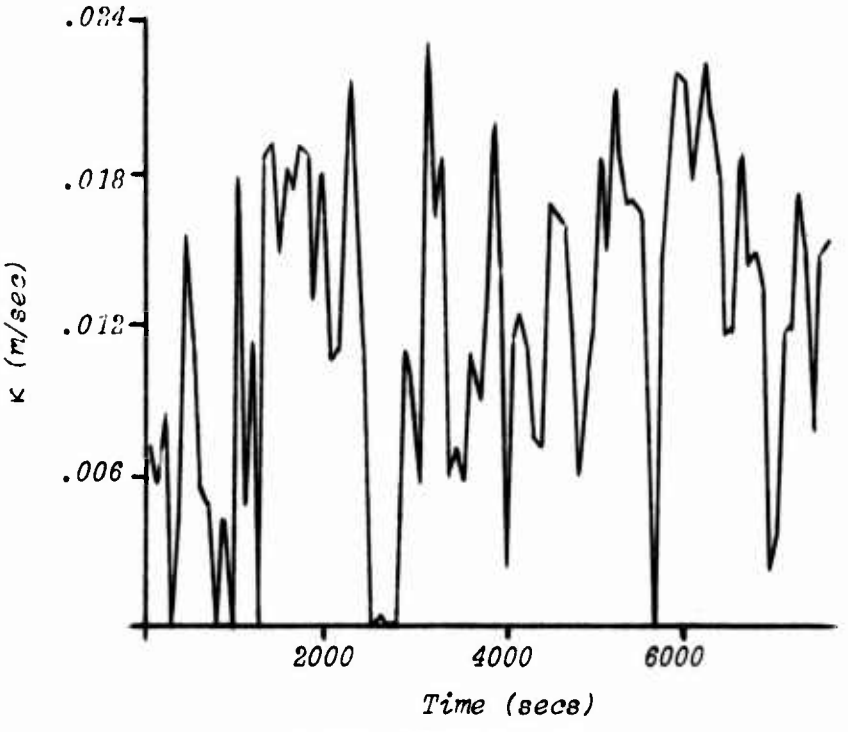
Figure 4.9 Range Rate Error, $\tilde{\dot{\rho}}$, and State Noise Standard Deviation, κ, for $S_0 = 1\ (m/sec)^4$

of the h_0-curve and the r-curve reveals a slight trend to increasing values. This is not surprising since we know the actual error is increasing, (the bias error is obvious here in Figure 4.9a), and hence the Q-filter is simply trying to adapt to this. Further, one can expect that divergence of $\dot{x}$ will eventually occur as the Q-filter saturates, and no new information is added. Recall that this is due to the assumption that the error is a stationary random process, i.e., constant Q. An obvious remedy is to include a process (state) noise variance term in the Q-estimator, as suggested by equation *(3.6.11)*. Another alternative is to improve the assumed model so that the assumption of constant Q is more accurate. This approach is considered in the discussion on estimating model errors.

A notable aspect of the sequential Q-filter performance is the fact that the 1σ curve has been found not to decrease as rapidly as for the ordinary Kalman filter. Consideration of the constraint on the estimate, i.e., $\hat{q} \geq 0$, reminds one that the filter is minimum variance only in the *a posteriori* sense. Thus the variance cannot decrease any more rapidly due to this constraint on Q. Further, large values of r result in slow decrease in S. This behavior happens to hold an advantage. Since the attendant filter does not decrease as rapidly, the filter operates for a longer period of time before saturating. On the other hand, the estimates take longer to "settle down." Some idea of the effect of different *a priori* values of the error variance, S, may be gained from Figures 4.10 and 4.11. Figures 4.10a and b show the $\dot{x}$ error performance and estimates of Q for $S_0 = S(t_0) = .01(m/sec)^4$. The performance of the range rate and Q estimates appears to be about the same as those in Figures 4.9a and b, although now 64.2% of the error is within 1σ. Similarly, at the final time the computed standard deviation in the estimate of $\dot{x}$ is $\sqrt{S} = .026(m/sec)^2$.

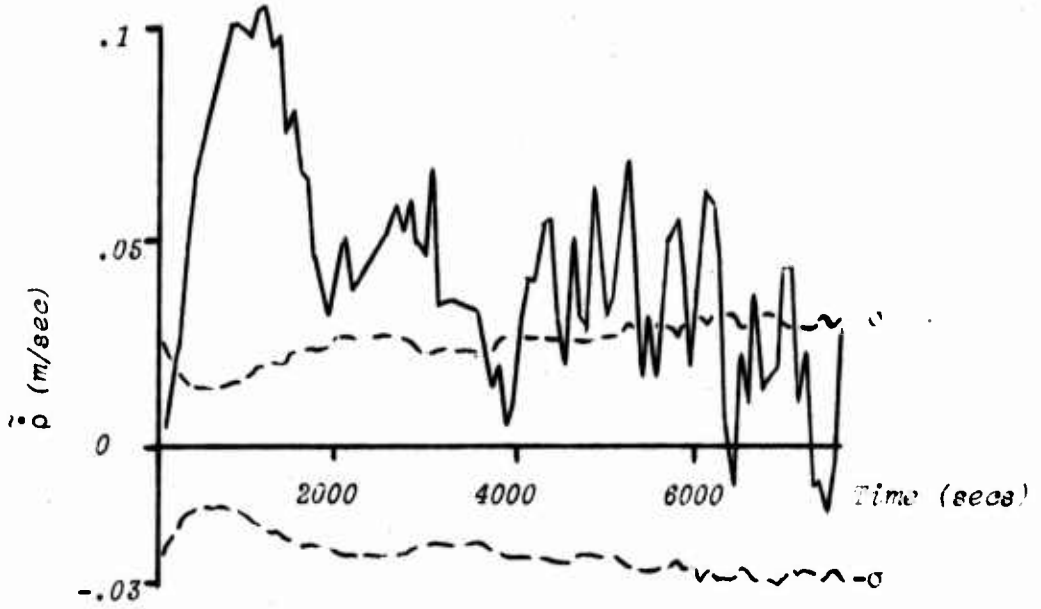


4.10a Range Rate Error, $\dot{\rho}$

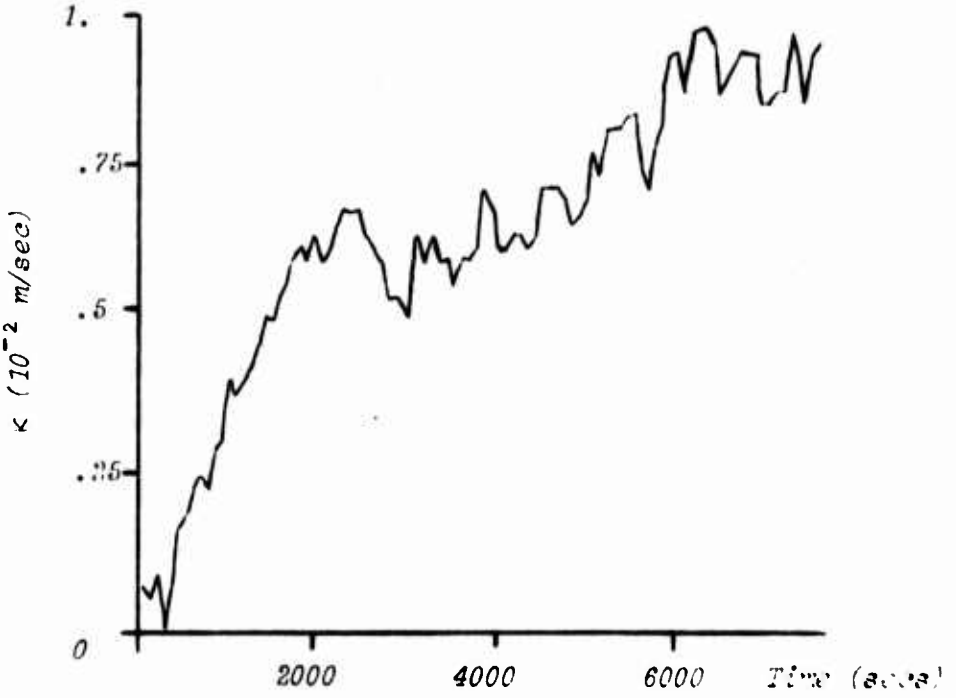


4.10b Standard Deviation, κ

Figure 4.10 Range Rate Error, $\tilde{\dot{\rho}}$, and State Noise Standard Deviation, κ, for $S_0 = 10^{-2}(m/sec)^4$



4.11a Range Rate Error, $\dot{\tilde{\rho}}$



4.11b Standard Deviation, κ

Figure 4.11 Range Rate Error, $\dot{\tilde{\rho}}$, and State Noise Standard Deviation, κ, for $S_0 = 10^{-4}(m/sec)^4$

Figures 4.11a and b show the $\dot{\rho}$ error and Q estimates for $S_0 = 10^{-4}(m/sec)^4$. Here the range rate estimate is not very good. A large bias appears early, although this decreases as larger Q-estimates are obtained (Figure 4.11b). Still, only 30.08% of the estimates lie under the 1σ curve. At the end of the simulation $\sqrt{S} = .983 \times 10^{-2}(m/sec)^2$.

A number of runs were made to investigate the effects of different *a priori* values of T. The results indicated that the *a priori* value has little effect upon both the estimates of T and Q. Using the algorithm of equation *(3.6.35)*, the estimated values of T decreased monotonically from 21.94 (m/sec)^4 to .2571 (m/sec)^4. Presumably if more points were taken, this would decrease even further. The modification, given by equation *(3.6.36)* of using a limited batch of Δ-residuals was not investigated for this problem, but is treated in the next chapter.

Estimation of Model Errors. Suppose the assumption of a constant closure rate is modified so as to be a linear function of time. We thus assume the model error may be approximated by equation *(3.3.6)*. If we further assume $c_0 = 0$ and $c \equiv c_1$, then the range rate model becomes

$$\dot{\rho}_k = \dot{\rho}_{k-1} + c_{k-1}\,\Delta t$$

$$c_k = c_{k-1} + w'_{k-1} \qquad (4.6.6)$$

where c_k is unknown, and is to be estimated along with $\dot{\rho}_k$. The term w' is a process noise term added to c and represents fundamental uncertainties in knowledge of c. Its statistics are $E\{w'\} = 0$ and $E\{w'^2\} = q$. There are now two state variables, and the problem may be cast in the canonical form as

$$x_k = \begin{bmatrix} \dot{\rho} \\ c \end{bmatrix}_k = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{\rho} \\ c \end{bmatrix}_{k-1} + \begin{bmatrix} 0 \\ w' \end{bmatrix}_{k-1} = \Phi_{k,k-1}\, x_{k-1} + w_{k-1} \qquad (4.6.7)$$

In this recast form, the approach is to use the standard minimum variance filter with the corresponding H-matrix being

$$H = [1 \quad 0] \tag{4.6.8}$$

In light of this extended form, it is important to raise the question of observability. Consider the information matrix given by equation (4.6.31) with $j = 1$.

$$T_{k,1} = \sum_{i=1}^{k} \Phi_{i,k}^{T} H_i^{T} R_i^{-1} H_i \Phi_{ik} \tag{4.6.9}$$

Substituting for Φ, H, and R,

$$T_{k,1} = \sum_{i=1}^{k} \begin{bmatrix} 1 & 0 \\ (t_i - t_k) & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} (100) [1 \quad 0] \begin{bmatrix} 1 & (t_i - t_k) \\ 0 & 1 \end{bmatrix}$$

$$= 100k \sum_{i=1}^{k} \begin{bmatrix} 1 & (t_i - t_k) \\ (t_i - t_k) & (t_i - t_k)^2 \end{bmatrix} \tag{4.6.10}$$

Forming the quadratic form using an arbitrary 2-vector, z, yields

$$z^T T_{k,1} z = 100k \sum_{i=1}^{k} [z_1 + z_2(t_i - t_k)]^2 \tag{4.6.11}$$

which is clearly positive for all non-zero values of z and $k > 1$. Thus our assumed dynamics represent a completely observable system. Note that this would not be true were we interested only in c, since the observations do not relate directly to c. Rather, information about c is obtained only through its correlation with $\dot{\rho}$. Thus in estimating c it is the cross-correlation, $\sigma_{\dot{\rho}c}$ (normalized by the variances $\sigma_{\dot{\rho}}^2$ and R) which comprises the gain, K_c, for c,

$$K_c = \sigma_{\dot{\rho}c} / (\sigma_{\dot{\rho}}^2 + R) \tag{4.6.12}$$

We note in passing that the system is also completely controllable since

$$\Lambda_{k,0} = q \sum_{i=1}^{k} \begin{bmatrix} (t_k - t_i)^2 & (t_k - t_i) \\ (t_k - t_i) & 1 \end{bmatrix} > 0$$

which follows from (1.3.25).

In carrying out the simulations for this formulation, the *a priori* values of $c = 0$ and $E\{\tilde{c}^2\} = 1.0(m/sec^2)^2$ are used. Initial results with $Q = E\{ww^T\} = 0$ are shown in Figure 4.12a. The performance is such that as the error covariance decreases, so does the gain, and hence the constant c tends to a particular value. However, the true error actually changes in a non-linear manner so the bias error begins to dominate after about 800 seconds. (Only 26.82% of the 384 points are within 1σ of the expected value.) This, of course, is due to the fact that $Q = 0$. Now examining the 1σ-curve suggests that a reasonable 1σ steady-state value of .02 m/sec for $\dot{\rho}$ might be acceptable. The value of Q which yields this 1σ-value may be found using the covariance prediction and update equations. For the prediction,

$$\bar{P} = \Phi P \Phi^T + Q$$

or

$$\begin{bmatrix} \bar{p}_{11} & \bar{p}_{12} \\ \bar{p}_{21} & \bar{p}_{22} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \Delta t & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & q \end{bmatrix} \qquad (4.6.13)$$

Carrying out the multiplication yields

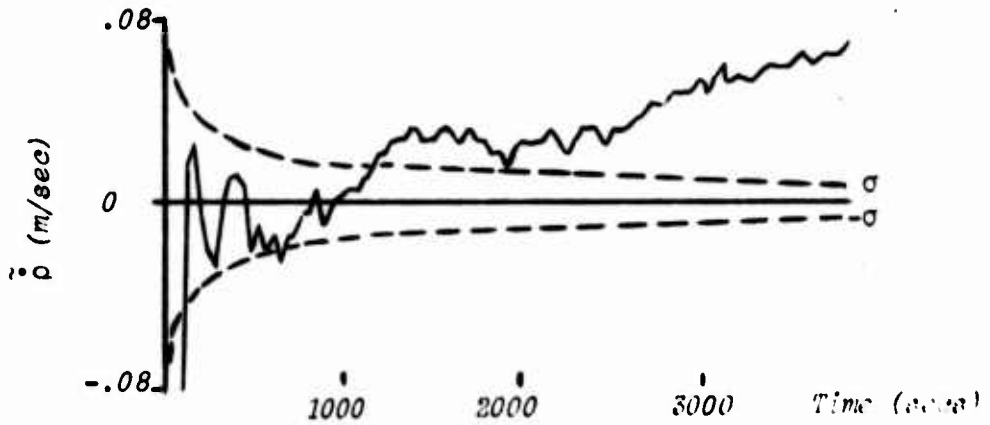
$$\bar{p}_{11} = p_{11} + 2p_{12}\,\Delta t + p_{22}\,\Delta t^2 \qquad (4.6.14a)$$

$$\bar{p}_{12} = p_{12} + p_{22}\,\Delta t \qquad (4.6.14b)$$

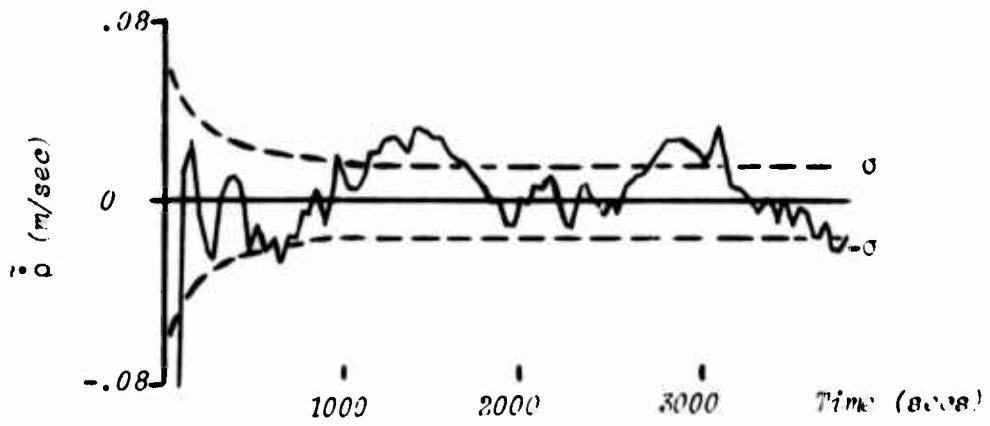
$$\bar{p}_{22} = p_{22} + q \qquad (4.6.14c)$$

where the symmetry of P is recognized as $p_{12} = p_{21}$.

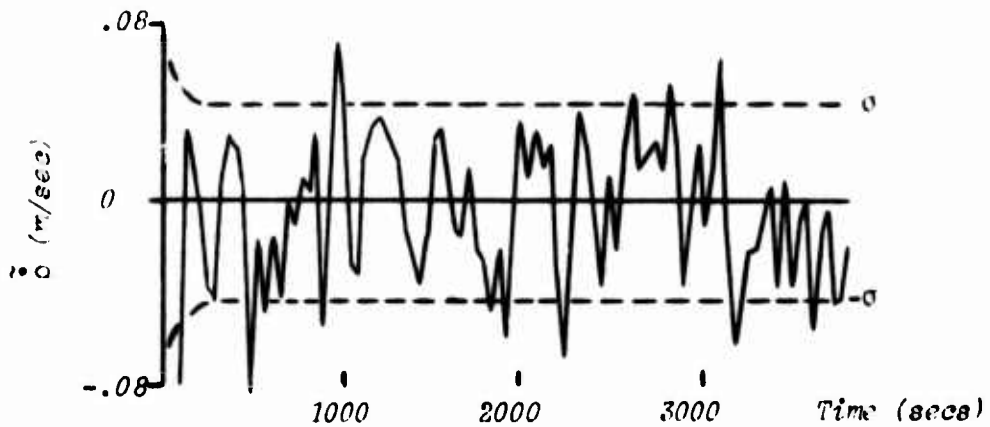
For the update equation,



4.12a $Q = 0\ (m/sec^2)^2$



4.12b $Q = 10^{-10}\ (m/sec^2)^2$



4.12c $Q = 10^{-7}(m/sec^2)^2$

Figure 4.12 Range Rate Error Linear Approximation for Various Constant Q Values

$$P = (1 - K H)\,\bar{P}$$

$$= \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} k_1 & 0 \\ k_2 & 0 \end{bmatrix} \right\} \begin{bmatrix} \bar{p}_{11} & \bar{p}_{12} \\ \bar{p}_{21} & \bar{p}_{22} \end{bmatrix}$$

or

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} (1-k_1)\,\bar{p}_{11} & (1-k_1)\,\bar{p}_{12} \\ -k_2\,\bar{p}_{11} + \bar{p}_{21} & -k_2\,\bar{p}_{12} + \bar{p}_{22} \end{bmatrix} \qquad (4.6.15)$$

Employing the definition of K, and again recognizing the symmetry of P,

$$p_{11} = \left(1 - \frac{\bar{p}_{11}}{\bar{p}_{11} + R}\right) \bar{p}_{11} = \frac{R\,\bar{p}_{11}}{\bar{p}_{11} + R} \qquad (4.6.16a)$$

$$p_{12} = \left(1 - \frac{\bar{p}_{11}}{\bar{p}_{11} + R}\right) \bar{p}_{12} = \frac{R\,\bar{p}_{12}}{\bar{p}_{11} + R} \qquad (4.6.16b)$$

$$p_{22} = -\frac{\bar{p}_{12}^{\,2}}{\bar{p}_{11} + R} + \bar{p}_{22} \qquad (4.6.16c)$$

Solving *(4.6.16a)* for $\bar{p}_{11}$ gives

$$\bar{p}_{11} = p_{11}\,R/(R - p_{11}) \qquad (4.6.17)$$

For steady state conditions, the p_{ij} in *(4.6.14)* equal the p_{ij} in *(4.6.16)*. Substituting *(4.6.14c)* into *(4.6.16c)* and solving for $\bar{p}_{12}$ yields

$$\bar{p}_{12} = \sqrt{q(\bar{p}_{11} + R)} \qquad (4.6.18)$$

Substituting this into *(4.6.16b)* gives the equation for p_{12} as

$$p_{12} = R\sqrt{q(\bar{p}_{11} + R)}\,/(\bar{p}_{11} + R) \qquad (4.6.19)$$

Next equate *(4.6.14b)* and *(4.6.18)*, and substitute *(4.6.19)* for p_{12} to obtain

$$p_{22} = \frac{1}{\Delta t}\sqrt{q(\bar{p}_{11} + R)}\left(1 - \frac{R}{\bar{p}_{11} + R}\right) \qquad (4.6.20)$$

Finally, use *(4.6.17)*, *(4.6.19)*, and *(4.6.20)* to solve *(4.6.14a)* for q. The result is

$$q = \frac{p_{11}^4}{\Delta t^2 (R - p_{11})(2R - p_{11})^2} \qquad (4.6.21)$$

Using the suggested steady state value of $\sigma = .02$ m/sec, then $p_{11} = \sigma^2 = 4 \times 10^{-4}(\text{m/sec})^2$ and, approximately,

$$q = .1 \times 10^{-9}(m/sec^2)^2$$

The values of Δt and R are taken from Table 4.1.

Using this value of q along with the other initial and *a priori* values ($e = 0$ and $E\{\tilde{e}^2\} = 1.0\ m^2/sec^4$) produces the results shown in Figure 4.12b. Here definite improvement is achieved as expected. 74.74% of the 384 points are within the steady state 1σ value of .0209 m/sec as opposed to the 26.82% associated with Figure 4.12a. Notable characteristics are the apparent oscillations of the error, and the fact that the majority of points lie above the time axis. Oscillatory error behavior is known to be characteristic of optimal approximations, e.g., (48, 75), and here optimality is clearly an objective. The latter characteristic implies, as in earlier runs, that the bias error is becoming apparent. An obvious suggestion is to increase q. The results for $q = .1 \times 10^{-6}(m/sec^2)^2$ are shown in Figure 4.12c. This arbitrary increase in q produces a steady state 1σ value of .053 m/sec, and bounds 83.33% of the error points. In light of these results, it appears that the optimum q is probably between $.1 \times 10^{-9}(m/sec^2)^2$ and $.1 \times 10^{-6}(m/sec^2)^2$.

To pursue the question of optimum Q, and also to investigate the sequential Q estimation algorithm further, the idea of estimating q arises. However, in applying the sequential algorithm strictly to the formulation given by *(4.6.7)* the observability problem is brought out again, this time with regard to estimating Q. In writing the information matrix, Γ^Q,

corresponding to the estimation of Q one finds that $T^Q = 0$. Because of the way the problem is formulated, the equation for $\bar{r}^2$ does not contain c explicitly. Thus in forming T^Q, the measurement mapping term is zero for all time, and hence so is T^Q. In other words, the term corresponding to H in equation *(1.3.21)* is zero. Hence, the Q corresponding to c is not observable.

Consider the formulation of the original problem as a differential equation. In vector form

$$\dot{x} = A\,x + B\,u$$

or

$$\begin{bmatrix}\ddot{\rho}\\ \dot{c}\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}\dot{\rho}\\ c\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix}u \qquad (4.6.22)$$

where the identifications are obvious. The term, u, is a process noise term with the statistics $E\{u\} = 0$ and $E\{u^2\} = q\,\delta(t-\tau)$. Now the solution to this differential form is clearly not *(4.6.7)*, but rather

$$\begin{bmatrix}\dot{\rho}\\ c\end{bmatrix}_k = \begin{bmatrix}1 & (t_k - t_{k-1})\\ 0 & 1\end{bmatrix}\begin{bmatrix}\dot{\rho}\\ c\end{bmatrix}_{k-1} + \int_{t_{k-1}}^{t_k}\begin{bmatrix}1 & (t_k-\tau)\\ 0 & 1\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix}u\,d\,\tau \qquad (4.6.23)$$

Taking the approach of equation *(1.2.30)*, where u is assumed constant over $[t_k,\ t_{k-1}]$, yields

$$x_k = \Phi_{k,k-1}\,x_k + \Gamma_{k,k-1}\,w_{k-1} \qquad (4.6.24)$$

where

$$\Gamma_{k,k-1} = \begin{bmatrix}\Delta t & \frac{\Delta t^2}{2}\\ 0 & \Delta t\end{bmatrix},\quad w_{k-1} = \begin{bmatrix}0\\ u_{k-1}\end{bmatrix} \qquad (4.6.25)$$

It is important to note that one form, *(4.6.7)* or *(4.6.24)*, is not necessarily more correct than the other. In both cases uncertainty has been introduced *arbitrarily*; hence the measure of correctness rests in the *choice* of the

model, and ultimately in the performance of the estimation process. For our purposes here *(4.6.24)* is indeed necessary since this appears to be the only way in which we can "get a handle" on Q.

To bring this approach to fruition, use *(4.6.24)* to form the error covariance prediction equation as

$$\bar{P} = \Phi P \Phi^T + \Gamma Q \Gamma$$

where

$$\Gamma Q \Gamma^T = \begin{bmatrix} \Delta t & \frac{\Delta t^2}{2} \\ 0 & \Delta t \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & q \end{bmatrix} \begin{bmatrix} \Delta t & 0 \\ \frac{\Delta t^2}{2} & \Delta t \end{bmatrix} = q \begin{bmatrix} \frac{\Delta t^4}{4} & \frac{\Delta t^3}{2} \\ \frac{\Delta t^3}{2} & \Delta t^2 \end{bmatrix} \qquad (4.6.26)$$

Now in forming the measurement equation to be used in the Q estimation we obtain

$$(\bar{r}^2 - H \Phi P \Phi^T H^T - R)_k = (H \Gamma Q \Gamma^T H^T + \Theta)_k$$

or

$$\bar{\Lambda}_k = \frac{\Delta t^4}{4} q_k + \theta_k \qquad (4.6.27)$$

Correspondingly, the information, Υ^Q, is no longer zero,

$$\Upsilon^Q_{k,1} = \frac{k T \Delta t^8}{16} > 0 \qquad (4.6.28)$$

and we have complete observability with respect to the $\bar{\Lambda}_k$.

While *(4.6.27)* is certainly a viable approach, it involves an approximation, namely that u is constant over the interval $[t_k, t_{k-1}]$. This assumption can be dispensed with by forming $\Gamma Q \Gamma^T$ in accordance with equation *(1.2.35a)*. Thus,

$$\Gamma Q \Gamma^T = \iint_{t_{k-1}}^{t_k} \Phi(t_k, s) B(s) q \delta(s-r) B^T(r) \Phi^T(t_k, r) \, dr \, ds$$

Carrying out the first integration,

$$\Gamma Q \Gamma^T = \int_{t_{k-1}}^{t_k} \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & q \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \Delta t & 1 \end{bmatrix} ds$$

or

$$\Gamma Q \Gamma^T = q \begin{bmatrix} \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} \\ \frac{\Delta t^2}{2} & \Delta t \end{bmatrix} \qquad (4.6.29)$$

The difference between *(4.6.29)* and *(4.6.26)* is approximately the factor, Δt. Forming the measurement equation we have

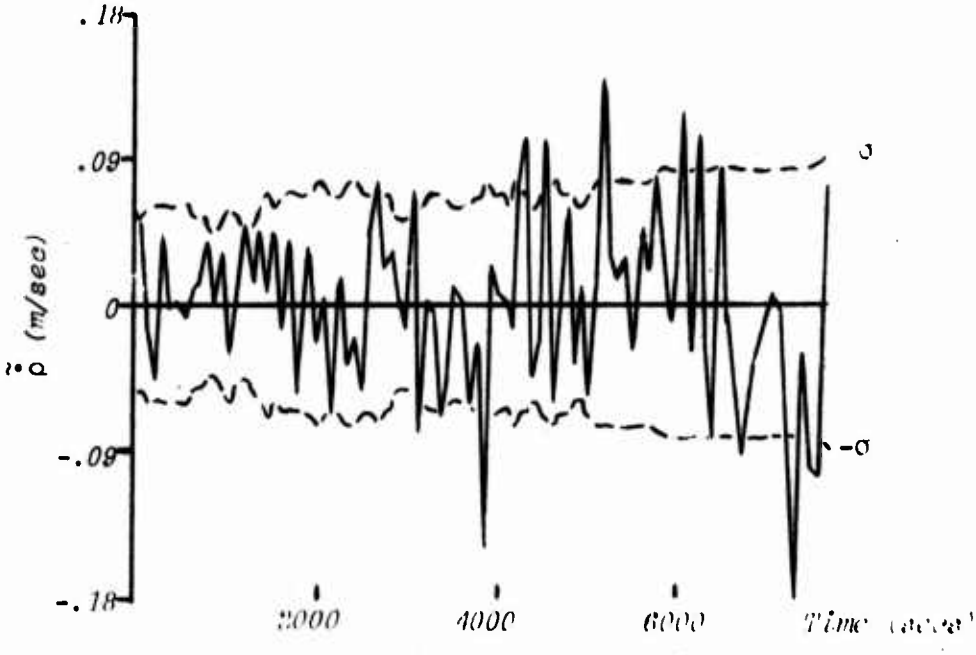
$$\bar{\Lambda}_k = \frac{\Delta t^3}{3} q_k + \theta_k \qquad (4.6.30)$$

and similarly the information

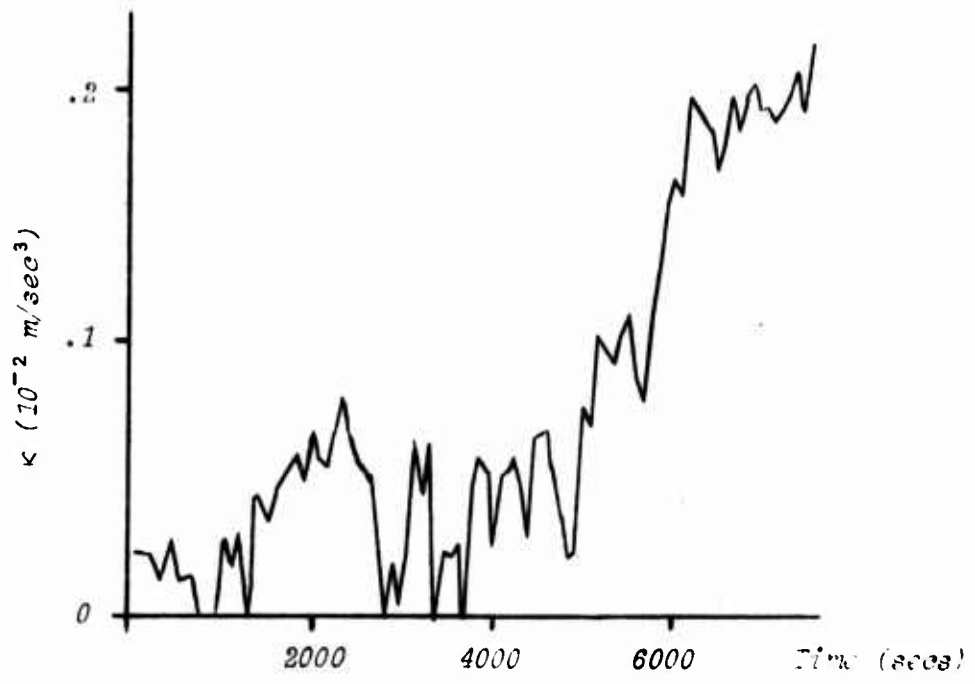
$$T^Q_{k,1} = \frac{k\ T\ \Delta t^6}{9} > 0 \qquad (4.6.31)$$

insures complete observability. Owing simply to the more rigorous approach, *(4.6.30)* will be used for the estimation of q.

Using this alternate formulation, Figures 4.13 through 4.15 show the estimation performance for the three *a priori* values, $S_0 = 10^{-8}$, 10^{-10}, and $10^{-12}(m/sec^2)^4$, respectively. The results indicate increasing estimation accuracy of the range rate with decreasing *a priori* values of S. In particular, for $S_0 = 10^{-8}(m/sec^2)^4$, 74.22% of the range rate error values are less than the 1σ curve (Figures4.13a). From Figure 4.13b the final value of q is approximately $4.5 \times 10^{-6}(m/sec^2)^2$. In Figure 4.14a, $S_0 = 10^{-10}(m/sec^2)^4$, 78.78% of the range rate error points are less than 1σ, and the maximum value of q (Figure 4.14b) is approximately $2. \times 10^{-8}(m/sec^2)^2$. For $S_0 = 10^{-12}$, 80.1% of the range rate errors are within 1σ of the expected zero mean (Figure 4.15a). The maximum q-value is approximately $1.7 \times 10^{-10}(m/sec^2)^2$ (Figure

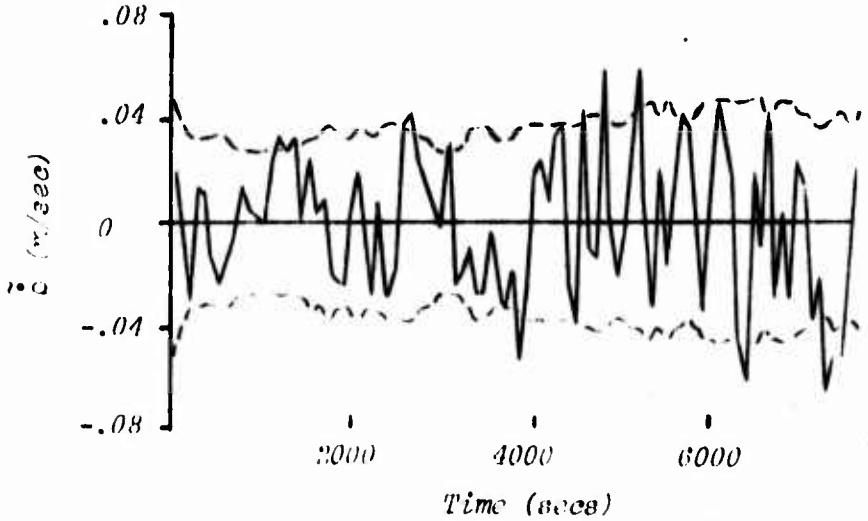


4.13a Range Rate Error, $\dot{\tilde{\rho}}$

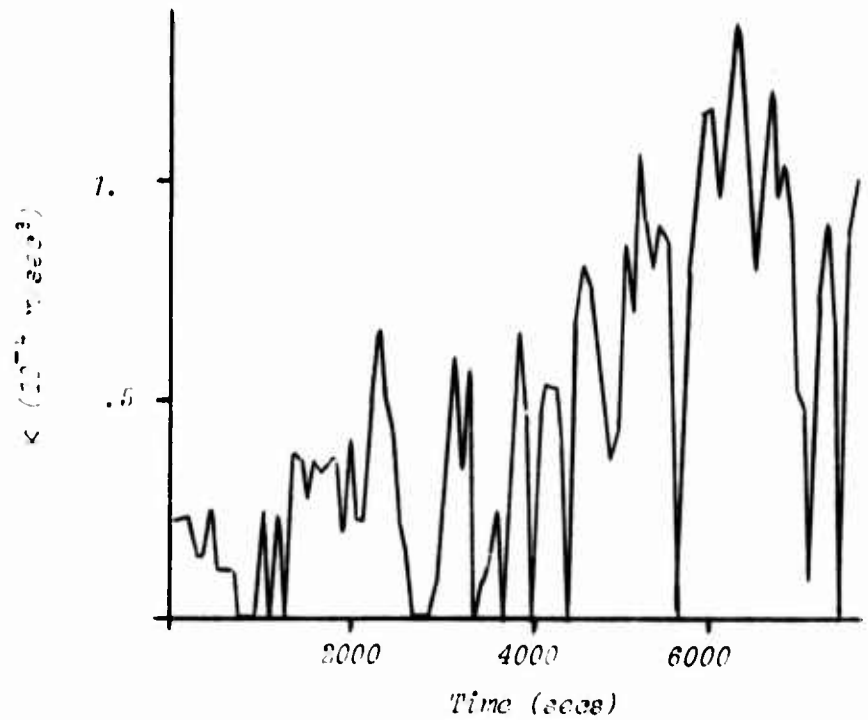


4.13b Standard Deviation, κ

Figure 4.13 Range Rate Error, $\dot{\tilde{\rho}}$, and State Noise Standard Deviation, κ, for Linear Error Approximation With $S_0 = 10^{-8}(m/sec^3)^4$

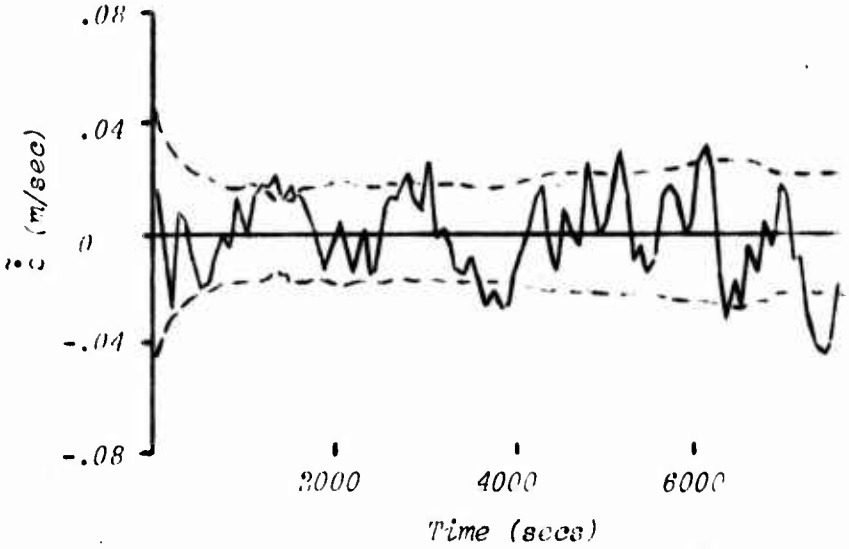


4.14a Range Rate Error, $\tilde{\dot{\rho}}$

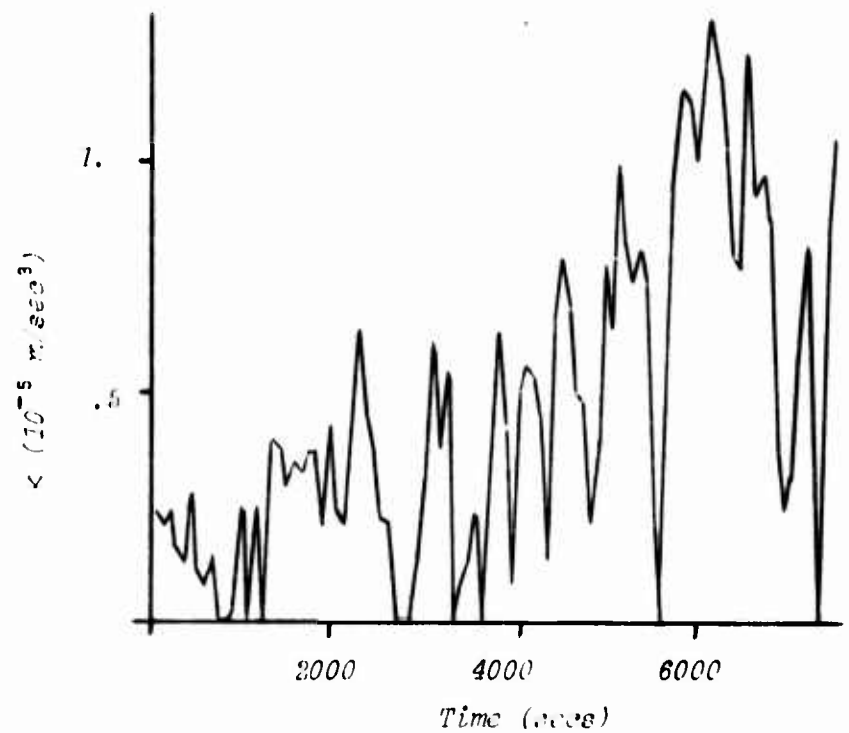


4.14b Standard Deviation, κ

Figure 4.14 Range Rate Error, $\tilde{\dot{\rho}}$, and State Noise Standard Deviation, κ, for Linear Error Approximation With $S_0 = 10^{-10}(m/sec^3)^4$



4.15a Range Rate Error, $\tilde{\dot{\rho}}$



4.15b Standard Deviation, κ

Figure 4.15 Range Rate Error, $\tilde{\dot{\rho}}$, and State Noise Standard Deviation, κ, for Linear Error Approximation With $S_0 = 10^{-12}$ $(m/sec^3)^4$

4.15a). While the actual magnitudes of the Q-estimates become smaller as the *a priori* S is decreased, within each case the sequence of Q values increases. This, of course, results in increasing the error variance of $\dot{\rho}$. As before, this may be attributed to the error growth, for while the error has been reduced, it has not been eliminated.

Note that for $S_0 = 10^{-8}(m/sec^2)^4$, comparatively large values of Q are obtained. This is due to the fact that initially large values of the gain are computed thus tending to make the values of $\dot{\rho}$ "follow" the observations. The result is that the erratic behavior of $\dot{\rho}$ from one observation to the next appears to be quite random. The sequential Q-estimator sees this and produces correspondingly large values of Q. As smaller values of S_0 are employed, this effect becomes less, thus allowing the values of $\dot{\rho}$ to become more stable. Figures 4.14a and 4.15a show this to some extent. The estimates of Q do not increase quite as rapidly; their magnitudes are less, and the values tend to reach their maximums sooner.

From these results it is apparent that various *a priori* values of S have definite effects upon the values of the Q-estimates. However, through reasonably careful selection of S_0, practical filter operation can be realized.

4.7 Summary and Conclusions

In this chapter the salient aspects of the estimation techniques found in Chapters 2 and 3 have been investigated, and the algorithms applied to the rendezvous problem formulated in Chapter 1. Explicit closed form expressions have been obtained, in whole or in part, for the age-weighting, ordinary and modified gain scaling, additive gain term, and limited memory filter algorithms. For these as well as all the other algorithms, numerical results have been obtained.

Certain equivalences were shown to exist in both the adaptive and non-adaptive forms of the age-weighting, modified gain scaling, additive gain term, and state noise covariance algorithms. While these, as well as the limited memory algorithm, were found to be acceptable estimators, the ordinary gain scaling technique (b = constant) was shown to be unacceptable, in both its adaptive and non-adaptive form.

The adaptive techniques, based on Jazwinski's maximum likelihood method as well as the sequential Q-estimator, are seen to be a step closer to accurate estimation. The techniques based on Jazwinski's approach have the particular advantages of being both simple and completely adaptive. The primary disadvantage is the fact that usually smoothing must be introduced in the form of a residual sample, thus destroying the optimality of the method.

The Kalman filter for sequentially estimating the state noise covariance appears to give slightly better performance than Jazwinski's adaptive estimator. Within the stated assumptions, the method is a constrained minimum variance estimator. However, the technique is more complex to implement. Further, the assumption of $E\{q\theta\} = 0$ destroys the optimality. If, in fact, Q does not represent a stationary process, then the assumption of constant Q introduces modeling error, thus requiring more accurate modeling or some technique to prevent saturation of the Q-filter.

The estimation of modeling errors has been shown to be particularly useful in that this approach attempts, in some way, to improve the dynamic model. Invariably the chosen structure is not sufficiently exact to allow the filter to operate to saturation. Thus, some model error compensation technique must still be employed, e.g., state noise covariance. In this regard, the sequential state noise covariance estimator has been shown to be effective.

The rendezvous problem investigated here involves a single state variable, and as such, the information and comparisons presented must be accepted in this light. When the algorithms are extended to state vectors of more than one variable, some of the algorithms lose their similarities. In particular, the modified gain scaling, additive gain term, and state noise covariance algorithms in both their adaptive and non-adaptive forms produce results which are generally different from one another. Further, Jazwinski's adaptive q-estimator requires a pseudo-inverse in computing a C-matrix. This disadvantage is not suffered by the sequential q-estimator. The limited memory filter is, in an *a posteriori* sense, an optimal estimator. However, applied to the vector case, it requires considerable numerical computation. To investigate these algorithms as applied to the vector case of more than one variable is an undertaking of considerably greater scope. Further, the insight provided by the closed-form expressions for many of the algorithms would be difficult, if not impractical, to obtain because of the greater problem complexity.

Based in general upon the results obtained in this chapter, the model error estimation technique and the sequential state noise covariance estimator have been selected for application to a multi-element state vector problem. The choice of this approach is due to the inherent potential for effective application to more complex problems. As indicated in Chapter 1, the problem is the orbit determination of a low thrust space vehicle which is subject to thrusting errors. This investigation is carried out in the next chapter.

Chapter 5

APPLICATIONS: LOW THRUST VEHICLE ORBIT DETERMINATION

5.1 Introduction

This chapter is concerned with the problem of estimating the state of a continuously thrusting, solar electric propulsion (SEP) space vehicle. The problem is compounded by the fact that the propulsion system is subject to certain mechanization and control errors. While errors arise from various other sources (49), for the SEP vehicle, the primary errors are due to anomalies in the propulsion system. Our aim here is to investigate the practicality of estimating not only the vehicle state, but the corresponding thrust acceleration errors as well.

A number of investigators have considered low thrust SEP missions. In (50), Rourke and Jordan investigated guidance and navigation approaches for two SEP interplanetary missions, although model errors were not estimated. Russell and Curkendall (36) obtained effective results by using piecewise constant functions to model acceleration errors. Tapley and Hagar investigated the estimation of acceleration errors (34) as well as inertial measuring unit errors (35) for an SEP vehicle. These approaches employed Kalman filtering utilizing the differential equation for P. Errors were successfully modeled as first and second order Gauss-Markov processes, although Earth rotational dynamics were not considered. Carpenter and Pitkin (37) investigated orbit determination for an SEP vehicle. Here the total thrust acceleration was assumed to be unknown, but approximated as the solution

to a set of uncoupled linear differential equations. However, the approach followed a least squares linearized simultaneous solution with no statistical measures being employed.

As indicated in the previous chapter, our approach here is to assume a structure for the acceleration errors. Then, employing the selected filtering equations, the parameters associated with the assumed model error structure, as well as the position and velocity of the vehicle are to be estimated.

In the following sections, the specific dynamics and estimation equations are developed. Five different model structures are investigated as to their ability to represent the acceleration errors. Each is developed separately, and numerical simulations are carried out to obtain comparative estimation performance. Corresponding to each of these structures, purely random errors are also assumed to be present. Associated covariances are either input, or estimated using the sequential Q-estimator.

5.2 Problem Description

In the problem considered, the motion of the solar electric spacecraft is assumed to be influenced by random errors in the thrust acceleration vector. The nominal SEP mission simulation is initiated at escape from the Earth's sphere of influence and terminates with a flyby of the asteroid Eros. Encounter with Eros occurs at a distance of 1.45 astronomical units (a.u.), 152 days after heliocentric injection.

If the only central force attraction considered is the sun, the equations of motion for the SEP spacecraft are

$$\dot{r} = v$$

$$\dot{v} = \frac{-\mu}{|r|^3} r + T \qquad (5.2.1)$$

where, as shown in Figure 5.1, r is a 3-vector of heliocentric position components, X, Y, Z; v is a 3-vector of heliocentric velocity components $\dot{X}$, $\dot{Y}$, $\dot{Z}$; $|r|$ is the magnitude of r; and μ is the gravitational parameter of the sun. T is the heliocentric thrust acceleration vector composed of the design thrust acceleration, T^*, as well as thrust acceleration errors, $n(t)$, from a number of sources (beam voltage and current, grid warpage, deadband control errors, etc.). The heliocentric components of T, $[T_X\ T_Y\ T_Z]^T$, may be expressed in a vehicle centered, orbit frame as $[T_x\ T_y\ T_z]^T$, where the two vectors are related by

$$T = \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = M \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$

where ψ is the heliocentric orientation angle (see Figure 5.1). The two reference frames are oriented such that the Z and z axes are parallel; the X and x axes form the angle, ψ, as do the Y and y axes, with

$$\cos\psi = X/(X^2 + Y^2)^{1/2}, \quad \sin\psi = Y/(X^2 + Y^2)^{1/2}$$

5.3 Acceleration Error Simulation

The SEP spacecraft is driven by an electric engine which in turn obtains its power from solar energy conversion devices, i.e., solar cells. While the actual solar flux density follows the inverse-square law, for outbound missions the actual available thruster power varies as $r^{1.7}$. This is due to improved efficiency of the solar cells at lower temperatures (50, 2). In addition to the thruster power variations, the propellant mass flow rate, $\dot{m}$, affects the magnitude of the thrust acceleration. In particular the magnitude, a, of the thrust acceleration, T, is given (50, 3) as

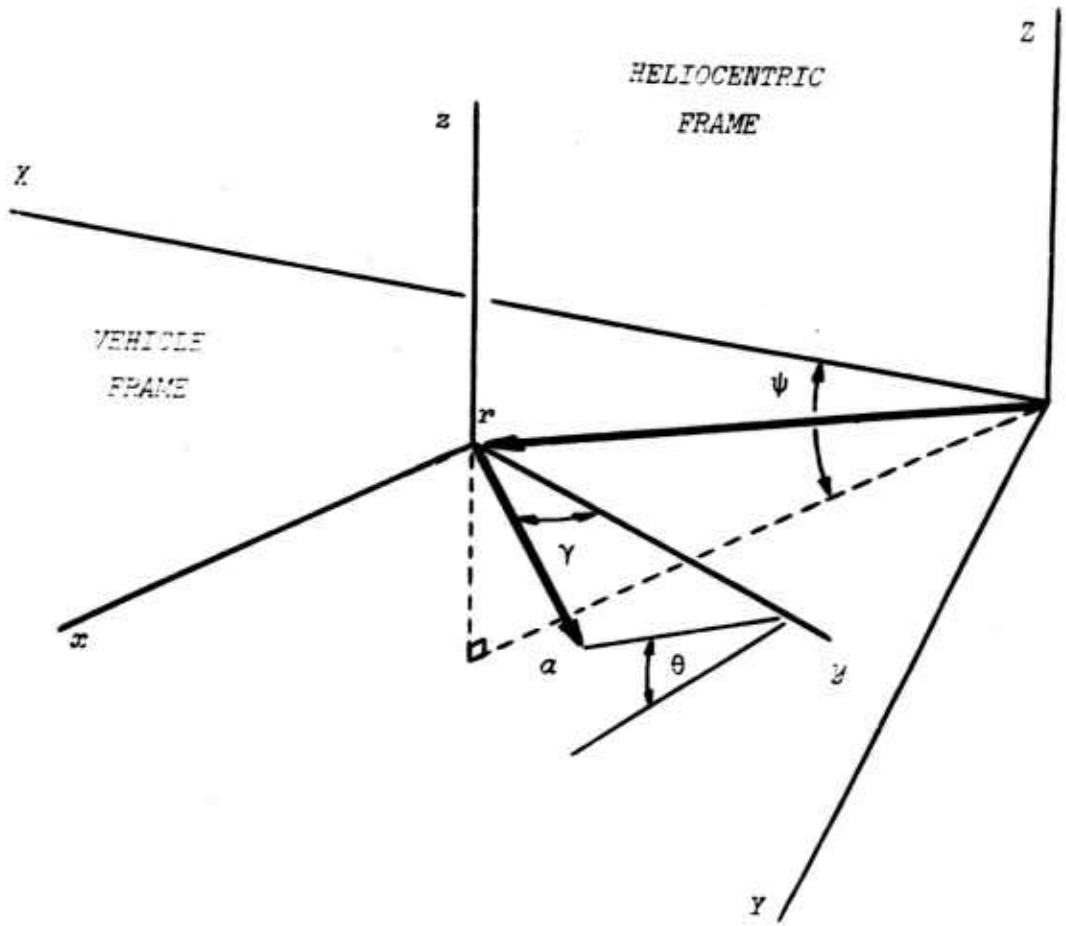


Figure 5.1 Reference Frames

$$a(t) = \frac{\dot{b}\, I_{sp}}{M\, g_0} \qquad (5.3.1)$$

where I_{sp} is the engine specific impulse, g_0 is sea level Earth gravitational acceleration, and M is the instantaneous spacecraft mass. These quantities in turn are functions of numerous propulsion system parameters, all with various uncertainties. Some of these have been mentioned -- beam voltage and current, grid warpage, and deadband control errors.

Regardless of the sources of error, their effect is to produce uncertainties in the thrust acceleration program. For the orbit determination function, we are interested in the errors at this total level, and it is at this level we propose to account for them. Now since we are not concerned here with the guidance problem, the design thrust program is arbitrary. Hence, for simulation purposes, a constant design thrust acceleration magnitude is selected. Further, the program is such that the nominal thrust acceleration vector, T^*, is oriented colinear with the y-axis of the orbital frame. The true thrust acceleration vector, T, can be resolved into its orbit frame components, T_x, T_y, and T_z, in terms of its magnitude, a, and the clock and cone angles, θ and γ, respectively (see Figure 5.1). Thus,

$$\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = a \begin{bmatrix} \sin\gamma & \cos\theta \\ \cos\gamma & \\ \sin\gamma & \sin\theta \end{bmatrix} \qquad (5.3.2)$$

and

$$a = a^* + \delta a \qquad (5.3.3)$$

where a^* is the nominal constant thrust acceleration magnitude and δa is the associated error. With zero error, the thrust acceleration is nominal. In this case,

$$T = T^A = a^A \begin{bmatrix} -\sin\psi \\ \cos\psi \\ 0 \end{bmatrix} \qquad (5.3.4)$$

In the presence of thrusting errors, γ, θ, and δa are non-zero thus producing an off-nominal thrust acceleration vector.

The acceleration error magnitude is simulated using

$$\delta a = \delta a_0 \sin \omega t + u_a \qquad (5.3.5)$$

where δa_0 and ω are constants and where u_a is a random variable with the statistics

$$E\{u_a\} = 0 \ , \quad E\{u_a^2\} = \sigma_a^2 \qquad (5.3.6)$$

In the error simulation the instantaneous values of the pointing angles, γ and θ, are assumed to be related as shown in Figure 5.2. This figure shows the x-z plane of the orbital frame. Assuming the cone angle, γ, is small, the radius of the circle is the maximum deviation, $\sin \bar{\gamma} \approx \bar{\gamma}$, of the normalized thrust vector* from its nominal position co-aligned with the y-axis. Next, the quantity

$$d = s(t - t_b) \qquad (5.3.7)$$

is the distance that the tip of the normalized thrust vector has moved since it last touched the boundary given by the circle of radius $\bar{\gamma}$. The rate, s, is simulated as a constant plus an additive noise component obtained from a normal distribution, $N(0, \sigma_s^2)$; t is the current mission time, and t_b is the time the boundary circle was last touched by the normalized thrust vector. The angle, ϕ, is obtained by sampling from a uniform distribution, $U(0, \sigma_\phi^2)$. With this information the instantaneous values of γ and θ can be obtained at each point in time. To clarify the process, the simulation logic is given

---

* That is, a unit vector in the direction of the thrust vector.

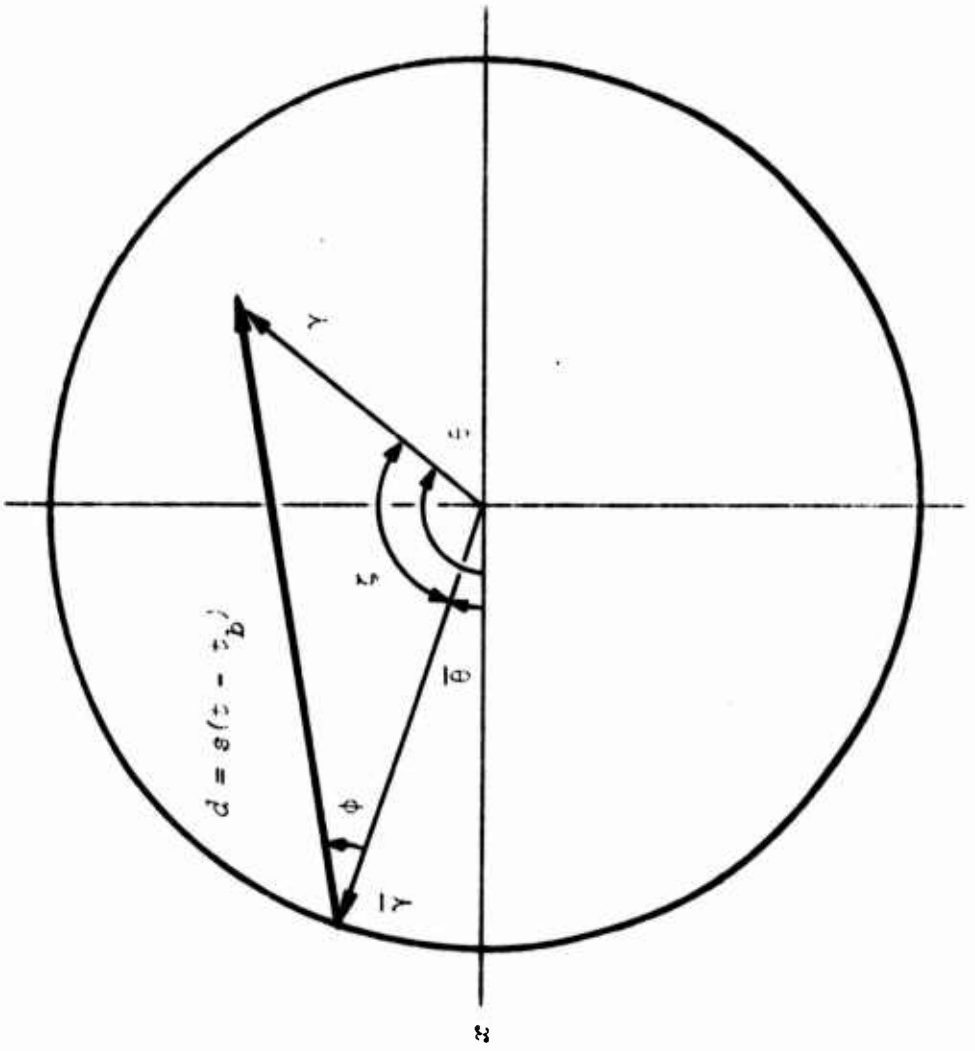


Figure 5.2 $\xi - \eta$ Acceleration Error Components

in Figure 5.3. Assume the initial values of $\bar{\gamma}$, ϕ, s, t_b, $\bar{\theta}$, and σ_s^2 and σ_ϕ^2 are given. To determine the values of γ and θ at t_k, the distance, d_k, is first computed. Using the law of cosines, the angles γ_k and ζ_k are determined. Then θ_k is computed by adding or subtracting ζ_k according to the sign of ϕ. Next a test is made to determine if the computed γ_k lies within the boundary circle. If it does, then sample values from the normal distributions, $N(0, \sigma_\gamma^2)$ and $N(0, \sigma_\theta^2)$ are added to γ_k and θ_k respectively. If the test is not passed, then new values of $\bar{\theta}$, ϕ, s, and t_b are found. Normally distributed noise is then added to γ_k and θ_k as before. For each point in mission simulation time, the sequence is repeated to obtain appropriate simulation values for the pointing angles.

5.4 Observation Geometry and Equations

Two observation types are employed for estimating the state vector. These are the radar-measured range rate, $\dot{\rho}$, of the vehicle, and the angle, ξ, formed by the lines-of-sight to the Earth and to a specified navigation star (see Figure 5.4). In computing range-rate, the motion of the tracking station, due to Earth rotation* as well as orbital revolution, is taken into account.

Consider Figure 5.4. The range vector can be expressed as

$$\bar{\rho} = r - r_s = r - R_s - R \qquad (5.4.1)$$

where r is the heliocentric position vector of the vehicle; r_s is the heliocentric position vector of the tracking station; R_s is the geocentric position vector of the station; and R_e is the heliocentric position vector of the Earth. In the figure, the reference frame $X'Y'Z'$ is a geocentric frame

---

* The diurnal effects of station motion have been shown to provide a significant contribution in obtaining accurate orbit determination (51, 34).

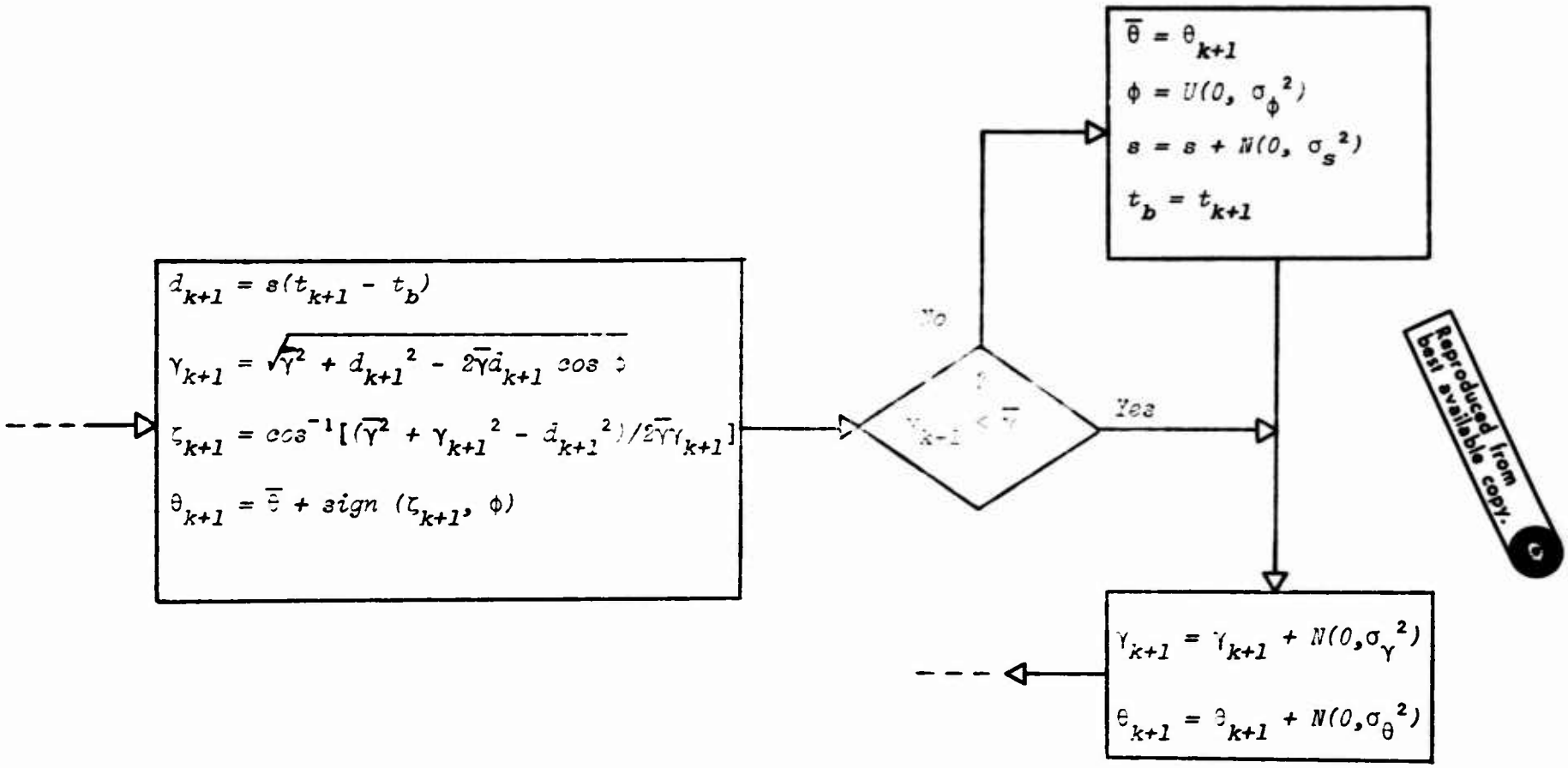


Figure 5.3 Pointing Angles Simulation

Reproduced from best available copy.

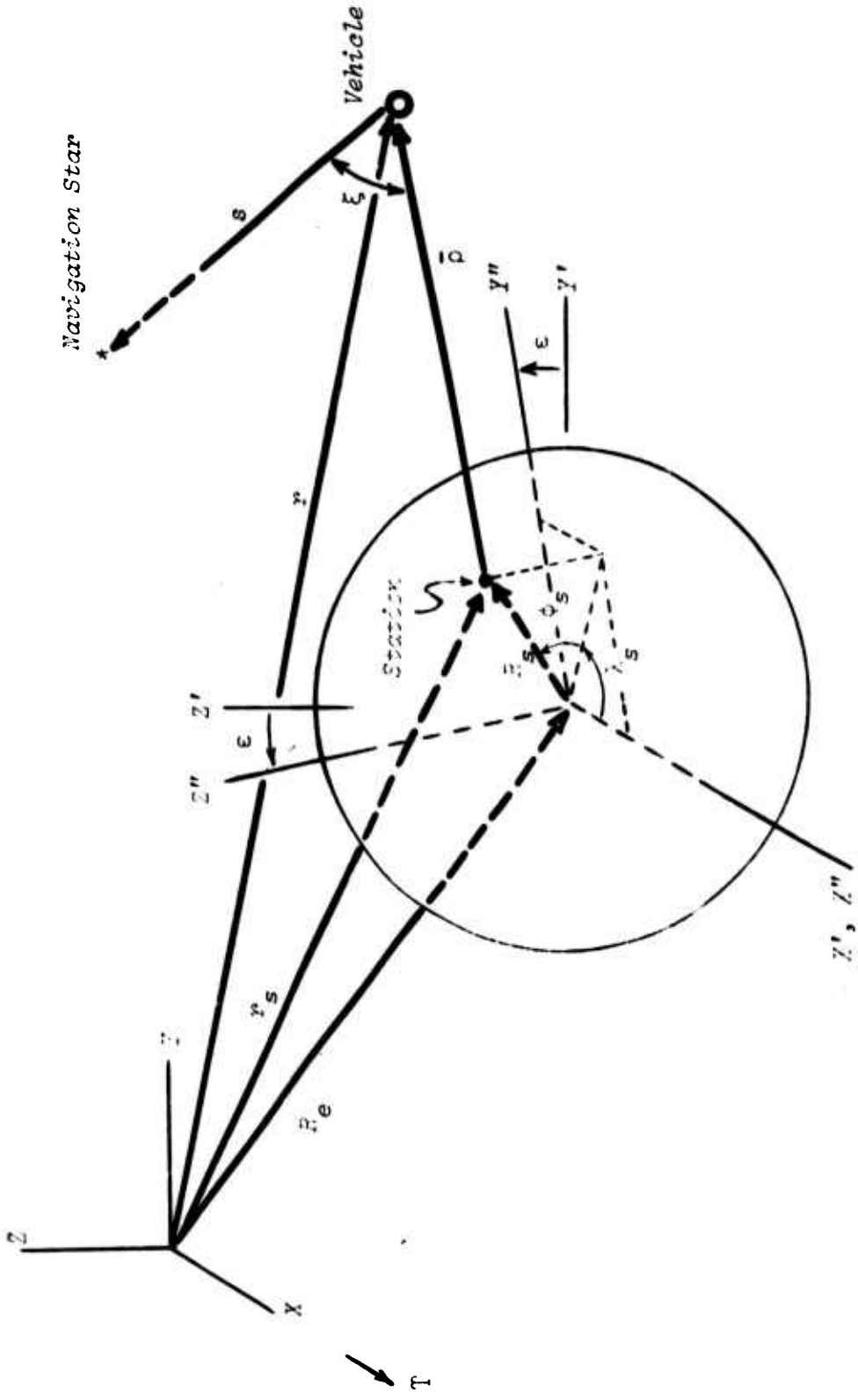


Figure 5.4 Observation Geometry

aligned with XYZ. The X and X' axes are assumed to be aligned, pointing to the First Point of Aires, T. $X''Y''Z''$ is a geocentric frame whose X''-Y'' plane contains the equator. Thus, $X''Y''Z''$ is rotated about X' through the angle, ε, the obliquity of the ecliptic ($\varepsilon \simeq 23\frac{1}{2}°$). The heliocentric components of R_e are simply

$$R_e = \begin{bmatrix} X_e \\ Y_e \\ Z_e \end{bmatrix} = |R_e| \begin{bmatrix} \cos \zeta \\ \sin \zeta \\ 0 \end{bmatrix} \tag{5.4.2}$$

where $|R_e|$ is the magnitude of R_e, and ζ is the Earth's heliocentric orientation angle. The heliocentric components of R_s are X_s, Y_s, and Z_s; they are given by

$$R_s = E\, R_s'' \tag{5.4.3}$$

where

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varepsilon & -\sin \varepsilon \\ 0 & \sin \varepsilon & \cos \varepsilon \end{bmatrix} \tag{5.4.4}$$

and

$$R_s'' = |R_s| \begin{bmatrix} \cos \phi_s & \cos \lambda_s \\ \cos \phi_s & \sin \lambda_s \\ \sin \phi_s & \end{bmatrix} \tag{5.4.5}$$

The angle, ϕ_s, is the geographic latitude of the tracking station, and λ_s is the right ascension of the station.

Now taking the derivative of *(5.4.1)* with respect to time gives

$$\dot{\rho} = \dot{r} - \dot{r}_s = \dot{r} - \dot{R}_s - \dot{R}_e$$

$$= \dot{r} - \dot{R}_s - \Omega_s \times R_s - \dot{R}_e - \omega_e \times R_e \tag{5.4.6}$$

where $(\dot{})$ is the relative derivative, Ω_s is the angular velocity of the station, and ω_e is the orbital velocity of the Earth. To simplify the computational process, the assumption is made that the Earth's orbit is circular and hence ω_e and R_e are constant. The angular velocity, Ω_s, is composed of the sidereal rotation rate, ω_s, and ω_e:

$$\begin{aligned}\Omega_s &= \omega_s + \omega_e \\ &= E\,\omega_s'' + \omega_e \end{aligned} \tag{5.4.7}$$

ω_s'' is aligned with Z'' and ω_e is aligned with Z. Thus

$$\dot{\bar{\rho}} = \dot{r} - \dot{r}_s$$

with

$$\dot{r}_s = (E\,\omega_s'' + \omega_e) \times R_s + \omega_e \times R_e \tag{5.4.8}$$

Multiplying out $E\,\omega_s''$ gives

$$\omega_s = \begin{bmatrix} 0 \\ -\omega_s'' \sin\varepsilon \\ \omega_s'' \cos\varepsilon \end{bmatrix} \tag{5.4.9}$$

Using (5.4.2) - (5.4.5), (5.4.7) and (5.4.9), equation (5.4.8) becomes

$$\dot{r}_s = \begin{bmatrix} (\omega_e - \omega_s'' \cos\varepsilon)\, X_s - \omega_s'' \sin\varepsilon\, Y_s \\ (\omega_e + \omega_s'' \cos\varepsilon)\, X_s \\ \omega_s'' \sin\varepsilon\, X_s \end{bmatrix} + R_e$$

or

$$\dot{r}_s = |R_s| \begin{bmatrix} (\omega_e \cos\varepsilon - \omega_s'') \cos\phi_s \sin\lambda_s - \omega_e \sin\varepsilon \sin\phi_s \\ (\omega_e + \omega_s'' \cos\varepsilon) \cos\phi_s \cos\lambda_s \\ \omega_s'' \sin\varepsilon \cos\phi_s \cos\lambda_s \end{bmatrix}$$

$$+ |R_e| \begin{bmatrix} -\omega_e \sin\zeta \\ \omega_e \cos\zeta \\ 0 \end{bmatrix} \tag{5.4.10}$$

with $\lambda_s = |\omega_s|\ t$ and $\zeta = |\omega_e|\ t$. Thus the desired range rate may be found as the magnitude, $\dot{\rho}$, of *(5.4.1)*.

The star-vehicle-Earth angle is measured on board the spacecraft. It is found simply as

$$\zeta = \cos^{-1}\left[\frac{\hat{s} \cdot (R_e - r)}{|R_e - r|}\right] \qquad (5.4.11)$$

For each simulated observation, random noise, v_i, is added to the deterministic value by sampling from a normal distribution. If the discrete observations at time t_i are specified generically as Ω, the associated statistics for $v_{\Omega i}$ are

$$E\{v_{\Omega i}\} = 0 \quad , \quad E\{v_{\Omega i}\ v_{\Omega j}\} = R_{\Omega i}\ \delta_{ij} \qquad (5.4.12)$$

5.5 Error Compensation Models

As indicated in section 5.2, the simulated thrust acceleration is composed of a nominal or programmed thrust acceleration, a, plus an error component, $m(t)$. The nominal dynamic model assumed for the estimation process is essentially the same as equation *(5.2.1)*. Of course, the accelerating error vector, $m(t)$, is unknown. We assume an approximation, $e(t)$, to $m(t)$, where $e(t)$ is modeled as a stochastic process with $\{e(t)\}$ satisfying one of several possible first or second order differential equations. The elements of $e(t)$ are the three orthogonal components of the acceleration error, and are expressed in the orbital frame.

While the general form given by equation *(3.4.6)* might be employed to obtain a single representation for the error, the selection and use of specific models offers the opportunity to investigate the effects of different structures, i.e., *structural sensitivity*. For this reason five different

models are employed: three basic forms, two of which have two separate configurations. These are described in the following paragraphs.

Model 0. In this model the acceleration error components are assumed to be represented as purely random process noise. Thus only a selected state noise covariance matrix is employed to maintain a positive definite error covariance, and hence filter operation, and to prevent divergence. In this case we define the estimated state vector to be $\chi^T = [r^T \; v^T]$, and the corresponding differential equations are

$$\dot{r} = v$$

$$\dot{v} = -\frac{\mu}{|r|^3} r + T^* + M u \qquad (5.5.1)$$

M is given by (5.2.2), and u is random state noise representing the acceleration errors. The *a priori* statistics for u are

$$E\{u\} = 0 \;, \quad E\{u_j(t)\, u_j(\tau)\} = q_j(t)\, \delta(t-\tau)$$

where $j = (x, y, z)$.

Model 1. The thrust acceleration error component along the orbit frame y-axis (in the direction of the nominal thrust) is approximated by a first order Markov process; the x- and z- components, in the orbit frame, are assumed to be purely random processes. Here the nominal differential equations are

$$\dot{r} = v$$

$$\dot{v} = -\frac{\mu}{|r|^3} r + T^* + M \begin{bmatrix} u_x \\ e \\ u_z \end{bmatrix} \qquad (5.5.2)$$

with the addition of one of the following two configurations:

a) $$\dot{e} = u_e \qquad (5.5.3)$$

b) $$\dot{e} = -\alpha e + u_e \qquad \dot{\alpha} = u_\alpha \qquad (5.5.4)$$

The random variables u_x, u_z, u_e, and u_α have the statistics

$$E\{u_l\} = 0 \ , \quad E\{u_l(t)\ u_l(\tau)\} = q_l(t)\ \delta(t-\tau)$$

where $l \in \{x, z, e, \alpha\}$. For each of the configurations a) or b) above, the corresponding state vectors to be estimated are

$$X = \begin{bmatrix} r \\ v \\ e \end{bmatrix} \qquad or \qquad X = \begin{bmatrix} r \\ v \\ e \\ \alpha \end{bmatrix} \tag{5.5.1}$$

respectively.

Model 2. The thrust acceleration error component along the orbit frame y-axis is approximated by a second order Markov process. The x- and z-components are again assumed to be purely random elements. The corresponding nominal differential equations are then

$$\dot{r} = v$$

$$\dot{v} = -\frac{\mu}{|r|^3}\, r + T^* + M \begin{bmatrix} u_x \\ e \\ u_z \end{bmatrix}$$

$$\dot{e} = g \tag{5.5.2}$$

plus one of the following configurations:

a) $$\dot{g} = u_g \tag{5.5.3}$$

b) $$\dot{g} = -\beta e + u_g$$

$$\dot{\beta} = u_\beta \tag{5.5.4}$$

The random variables u_x, u_z, u_g, and u_β have the *a priori* statistics

$$E\{u_l\} = 0 \ , \quad E\{u_l(t)\ u_l(\tau)\} = q_l(t)\ \delta(t-\tau)$$

where $l \in \{x, z, g, \beta\}$. For each of these above configurations the corresponding state vectors to be estimated are

$$\chi = \begin{bmatrix} r \\ v \\ e \\ g \end{bmatrix} \qquad or \qquad \chi = \begin{bmatrix} r \\ v \\ e \\ g \\ \beta \end{bmatrix}$$

respectively.

In the ensuing discussions the particular models and state vectors will be referred to often by their model number and letter configuration. Thus model 2b is that described by model 2, equation (5.5.8), and configuration b), equation (5.5.10).

Note that each of the above model forms seeks to approximate $m(t)$ in its own way. Model 0, of course, has a null structure as defined in section 3.3. Model 1a approximates $m(t)$ by a sequence of constants (Figure 5.5a); model 1b uses a sequence of exponentials (Figure 5.5b). Model 2a employs straight line segments (Figure 5.5c), and 2b approximates $m(t)$ by a sequence of arcs corresponding to the output of the simple harmonic oscillator.

For any of the assumed models, the state vector differential equation can be written in the general form

$$\dot{\chi}(t) = F(\chi, t) \qquad (5.5.11)$$

In view of (5.2.1), $F(\chi, t)$ is a nonlinear vector function; hence a suitable estimation procedure is extended form of Algorithm I. This is basically the same as Algorithm II, except that here the integral form for $\bar{\bar{P}}_k$ is replaced with the discrete form given by (1.2.17).

In forming $\bar{P}$, given by (1.2.17), the term $\Gamma Q \Gamma^T$ is required. Two methods for determining this are given by equations (1.2.30), (1.2.35a), and (1.2.35a). The latter form,

$$\Gamma Q \Gamma^T = \int_{t_{k-1}}^{t_k} \Phi \, B \, U \, B^T \, \Phi^T \, d\tau \qquad (1.2.35a)$$

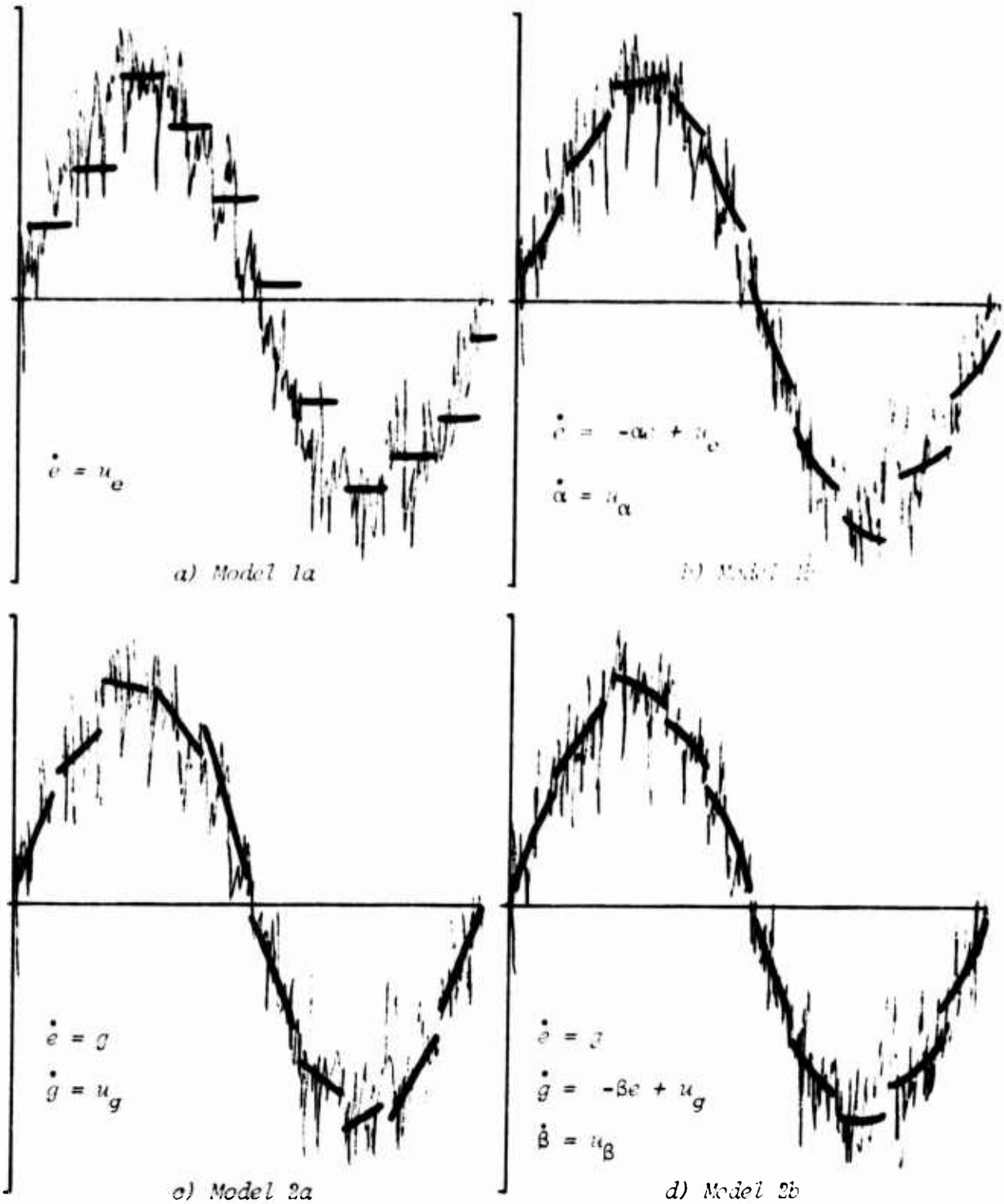


Figure 5.5 Approximation Models

is the more correct form, and is used exclusively when dealing with constant input values for $U\ \delta(t-\tau) = E\{u(t)\ u(\tau)\}$. In this case the elements of Q are the terms q_l where l is taken from one of the index sets associated with Model 0, 1, or 2. Numerical computation of $\Gamma Q \Gamma^T$ is performed as a simple quadrature based on the mean value theorem of integral calculus:

$$\Gamma Q \Gamma^T = (B_k\ U\ B_k^T + \Phi_{k,k-1}\ B_{k-1}\ U\ B_{k-1}^T\ \Phi_{k,k-1}^T)\ \frac{\Delta t}{2} \qquad (5.5.12)$$

For the cases where the sequential γ estimation algorithm is employed, equation *(1.2.33a)* is used with Γ given by *(1.2.30)*:

$$\Gamma = \int_{t_{k-1}}^{t_k} \Phi\ B\ d\tau \qquad (1.2.30)$$

This form is required because the Γ matrix itself is used. In this case we let $U = Q$, and the elements are again the terms q_l with l taken from the appropriate index set. The numerical computation of Γ is carried in a manner similar to the above, again using the mean value theorem:

$$\Gamma = (B_k + \Phi_{k,k-1}\ B_{k-1})\ \frac{\Delta t}{2} \qquad (5.5.13)$$

The elements of B for all models are given in Appendix B.

5.6 Simulation Process

Figure 5.6 is a functional flow diagram of the simulation process. It represents the logic followed by the simulation computer program. Input data is read from an appropriate input file, and various problem parameters and logic switches are initialized. Both the simulated and nominal trajectories are numerically integrated simultaneously through one observation interval. The simulated thrust acceleration is computed according to Figure 5.3, and the nominal thrust acceleration computed from one of the appropriate error models of equations *(5.5.1)* through *(5.5.10)*. In addition, the

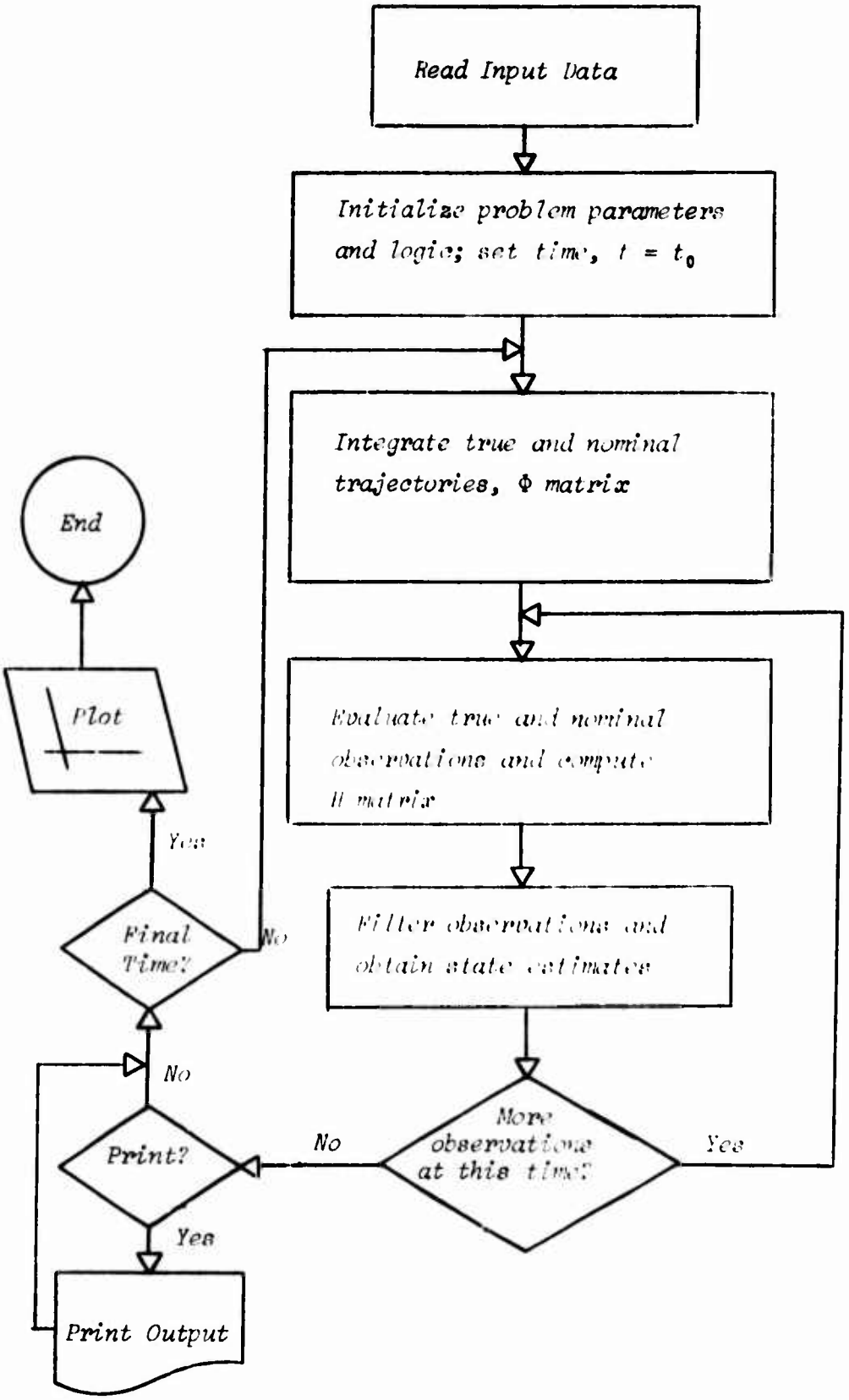


Figure 5.6 Simulation Logic Flow

differential equation for the state transition matrix,

$$\dot{\Phi}(t, t_j) = A(t)\ \Phi(t, t_j) \qquad (1.2.29)$$

is simultaneously integrated. As discussed in section 1.2, the matrix, A, is the partial derivative of F with respect to the state, χ. The elements of A for each model are given in Appendix B.

Next the first simulated observation is computed with a random noise component being added. (Observations are processed one at a time; hence, we have the computational advantage of scalar observation.) The corresponding nominal observation value is determined, and the observation mapping matrix, H, evaluated. The elements of H for each observation are given in Appendix C.

The estimation equations are employed to obtain the state estimate, $\hat{\chi}$. If any further observations are to be processed at this time point, the appropriate logic is repeated as shown in the figure.

A test is made to determine if print output is required, and if so, the appropriate information is written. A test is also made to determine if the final simulation time has been reached. If it has not, the process of integration, observation, and estimation is repeated for the next and succeeding observation intervals until the final time is reached. Finally, data plotting is accomplished as determined from input data and the simulation run terminates.

5.7 Numerical Results

In performing the various numerical simulations, a common set of basic problem data is consistently used. This approach provides a common basis for evaluating the estimation performance of each of the approximating models. This set of data is given in Table 5.1. The initial conditions, position and velocity, are the same for both the simulated and nominal trajectories. The

Initial Conditions		Initial Uncertainties	Thrust Acceleration Error Simulation
X	$.1505 \times 10^9$ km	10^4 km	$\delta a_0 = 6\sigma\, a^* = 1.8 \times 10^{-2}$ mm/sec^2
Y	0. km	10^4 km	$\sigma_a = .6\sigma\, a^* = 1.8 \times 10^{-3}$ mm/sec^2
Z	-6378. km	10^4 km	$s = .6 \times 10^{-6}$ rad/sec, $\sigma_s = .24 \times 10^{-7}$ rad/sec
$\dot{X}$	0. m/sec	10^3 m/sec	$\sigma_z = \sigma_\theta = .001$ rad, $\sigma_\varphi = .91$ rad
$\dot{Y}$	31,743 m/sec	10^3 m/sec	$\bar{\gamma} = .01745$ rad, $\omega = .12 \times 10^{-5}$ Hz
$\dot{Z}$	73.8 m/sec	10^3 m/sec	Period of $\delta a = 2\pi/\omega = 9.65$ days

Observation Variances	Nominal Thrust Acceleration
$R_\rho = (.5$ mm/sec$)^2$ @ 1 min intervals	$\ddot{x} = \ddot{z} = 0$ mm/sec^2
$R_\xi = (7$ arsec$)^2$ @ 10 min intervals	$\ddot{y} = a^* = .3$ mm/sec^2
Observation interval, $\Delta t = 50$ min	(Orbital Reference)

Tracking Station	Longitude	Latitude	Constants
Woomera (JPL-41)	136° 53' 14"	-31° 22' 55"	$\mu_{sun} = .1327 \times 10^{12}$ km^3/sec^2
Goldstone (JPL-14)	243° 06' 37"	35° 25' 33"	$\omega_e = .729 \times 10^{-4}$ rad/sec
Madrid (JPL-61)	355° 45' 3"	40° 25' 44"	$\omega_s'' = .727 \times 10^{-4}$ rad/sec

Navigation star unit vector elements: $\hat{s} = .5538, .7384, .3846$

Table 5.1 Nominal Mission and Simulation Parameters

Reproduced from best available copy.

observation interval is a constant 50 minutes for each observation type; hence the corresponding variances are scaled to agree with this sampling rate. The initial geometry is such that the Greenwich Meridian intersects the X-axis at midnight, GMT*, forming the spacecraft initial subpoint location.

In order to gain an idea of the acceleration error components, the plots shown in Figures 5.7 and 5.8 were obtained. Figure 5.7 shows the error components along each of the orbital frame axes as functions of time. Note the periodicity of the y-component and the irregularity of the x- and z-components. Further, the y-component is approximately three times as great as the others. Figure 5.8 shows the trace of the thrust acceleration vector projected on the orbital frame x-z plane. Note the approximately circular bound of radius $\simeq .0055\ nm/sec^2$. The errors are also seen to be somewhat concentrated in the first quadrant, and rather less dense in the fourth quadrant. As the simulation time continues beyond the 35 days shown here, one can expect that these errors would be more uniformly distributed within the full region.

In the following paragraphs numerical results are presented showing estimation performance for the various models and approaches. In presenting the data, two quantities are plotted which represent a figure of merit of performance. These are the Euclidean norms of the error components of position and velocity, and the square root of the trace of the appropriate covariance submatrix elements. The ensuing discussions will refer to these quantities as RSS (root-sum-square) and RTC (root-trace-covariance), respectively.

---

* GMT = Greenwich Mean Time; see, for example, (9).

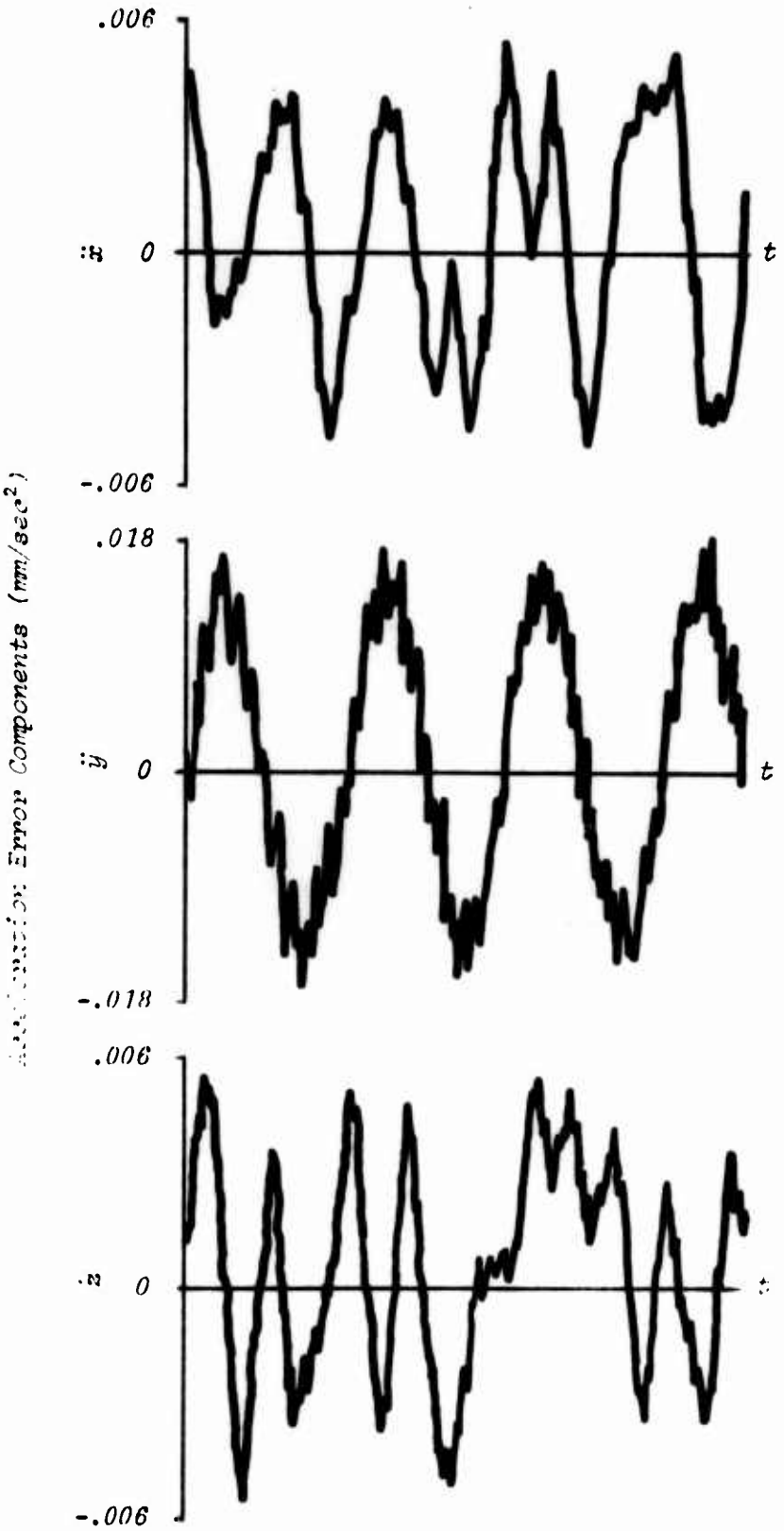


Figure 5.7 Acceleration Error Components
(t = 35 days)

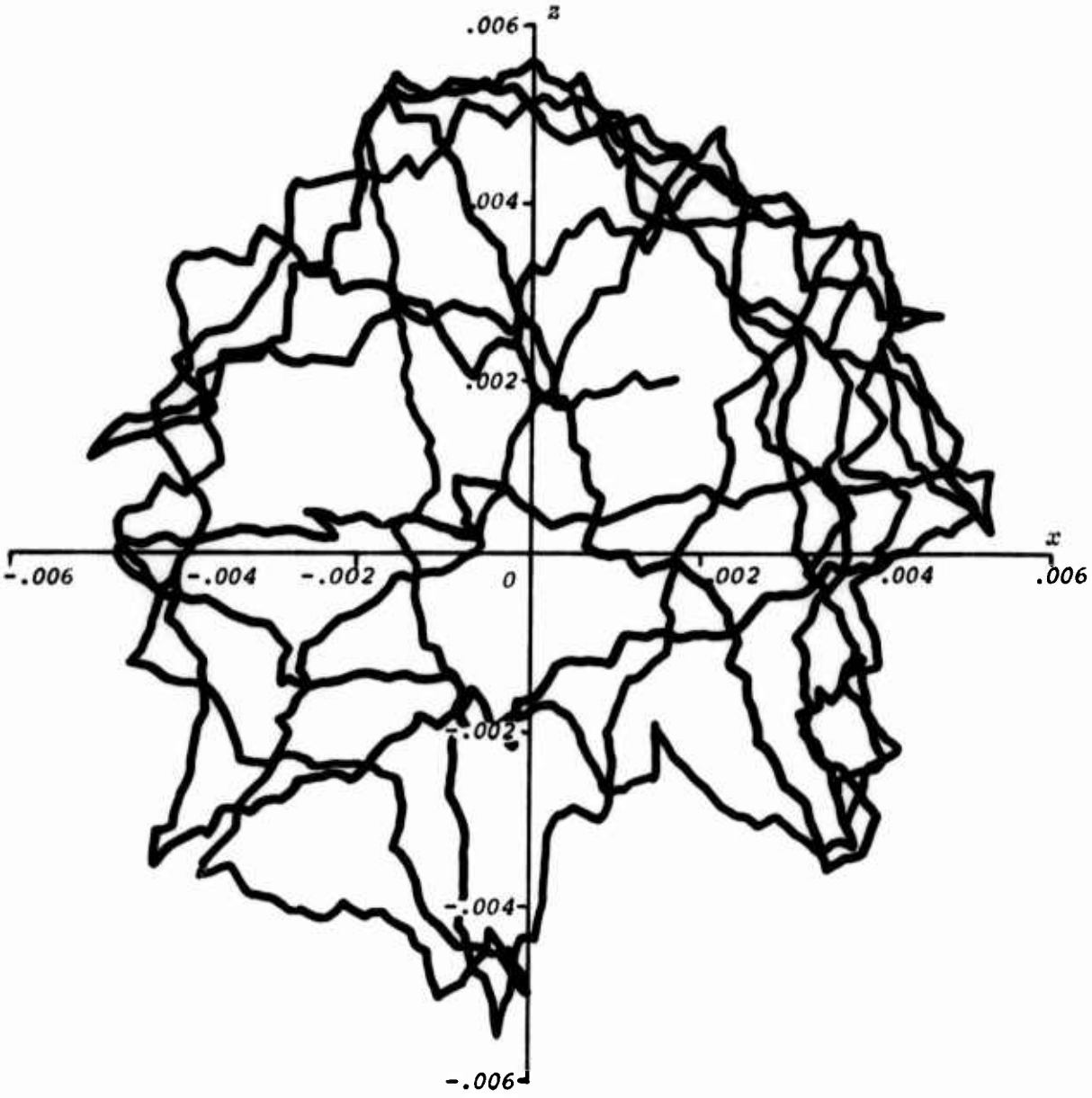


Figure 5.8 Acceleration Errors in x-z Plane (mm/sec^2)

Nominal Cases With and Without Model Error. Figures 5.9 through 5.12 show the position and velocity error norm, RSS, for the nominal data of Table 5.1, both with and without thrust acceleration modeling error, respectively. Clearly the estimation performance in Figures 5.9 and 5.10 is erratic and unacceptable, producing estimation errors as high as 20,000 km and 18 m/sec. The rapid decrease in the error covariances is illustrated by the short dashed line, the RTC, near the origin. Figures 5.11 and 5.12 illustrate the case of perfect modeling. Here both the actual and assumed thrust accelerations are of constant magnitude, always directed along the orbital frame v-axis. Two RSS curves are shown in each figure, corresponding to two different values of the observation error variance for the star-Earth angle, ξ. In both cases the RSS values become substantially less than 2 km in position and .004 m/sec in velocity as the estimation process proceeds.

Perhaps more interesting is the sensitivity, shown by these curves, of the estimation process to differences in the observation error variance. The solid line in Figures 5.11 and 5.12 represents the RSS for the nominal (Table 5.1) value of $R_\xi = (7\ arcsec)^2$. The dashed line corresponds to an increase of 3 arcsec resulting in $R_\xi = (10\ arcsec)^2$. In both cases, of course, each value is used for both the true (simulated) and assumed (nominal) value. This change in the value of R_ξ illustrates a somewhat surprising sensitivity of the estimation accuracy to this particular measurement type. For the increased R_ξ the position RSS is more than twice that for the nominal value. Clearly similar behavior is seen for the corresponding velocity RSS.

The sensitivity of the estimation process to R_ξ, and the problems of actually developing equipment capable of such high accuracies prompts the question of necessity. Are ξ measurements indeed necessary? To answer this, consider the information obtained from a single range-rate observation. From

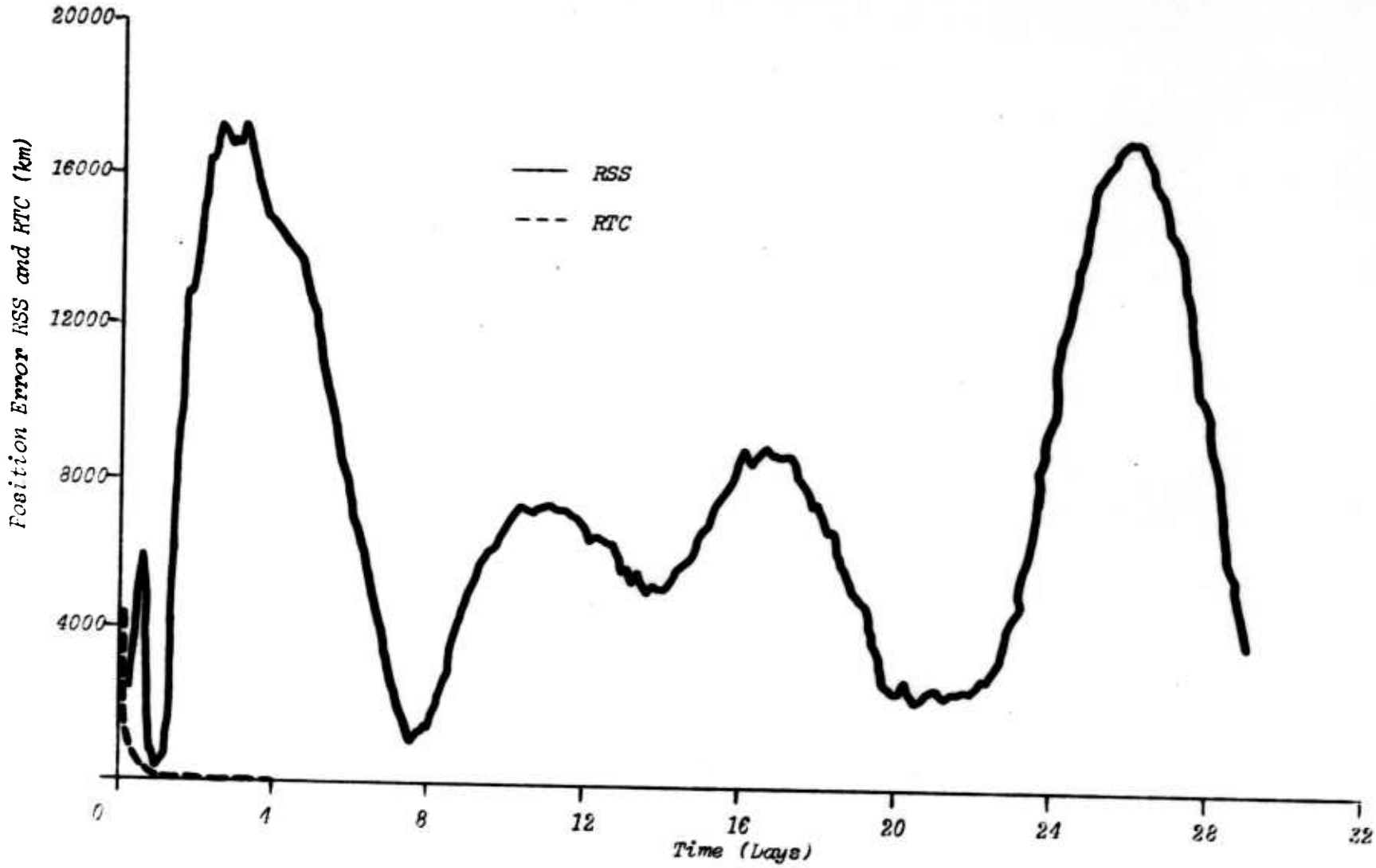


Figure 5.9 Position Error, No Model Error Compensation

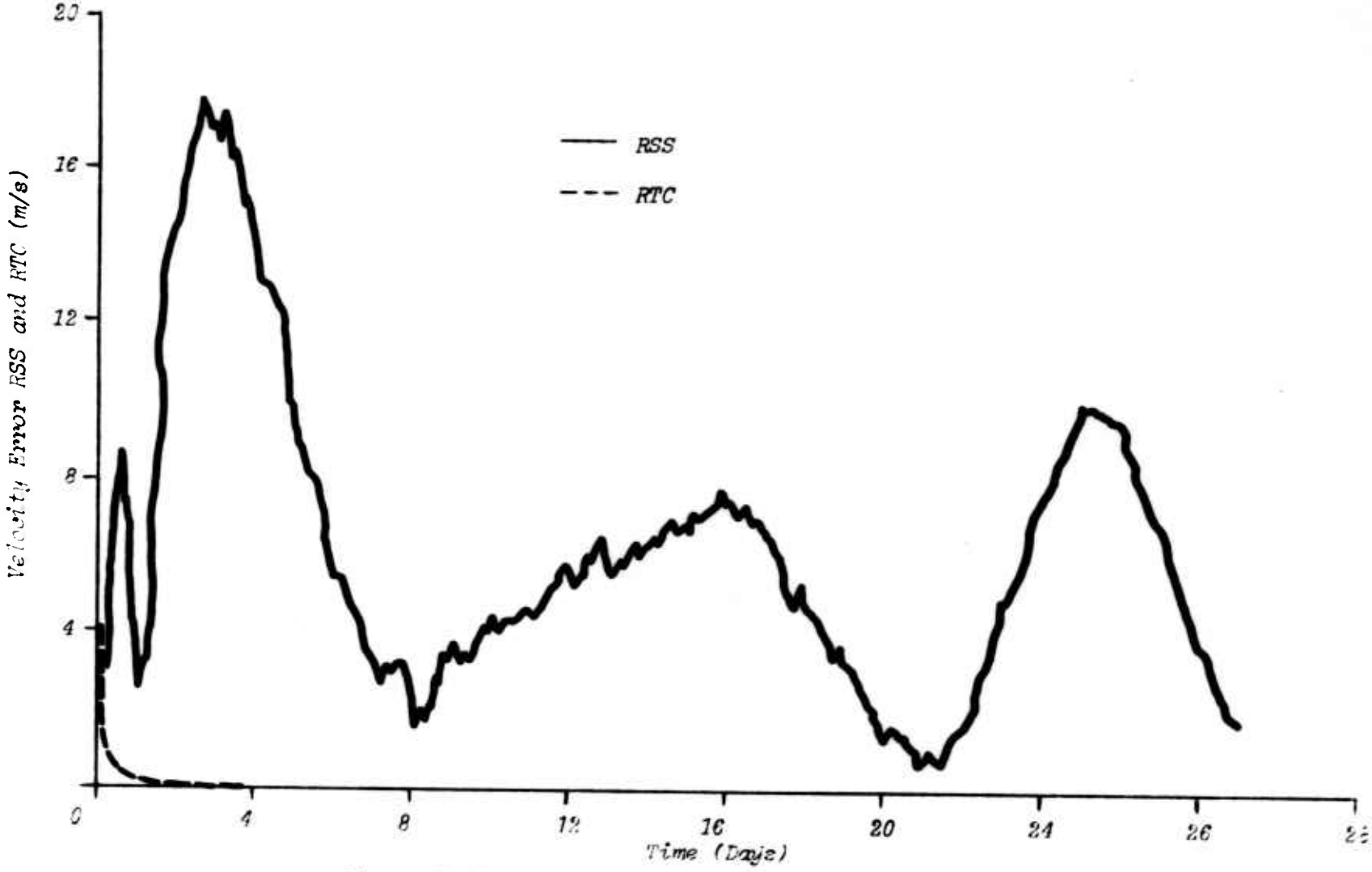


Figure 5.10 Velocity Error, No Model Error Compensation

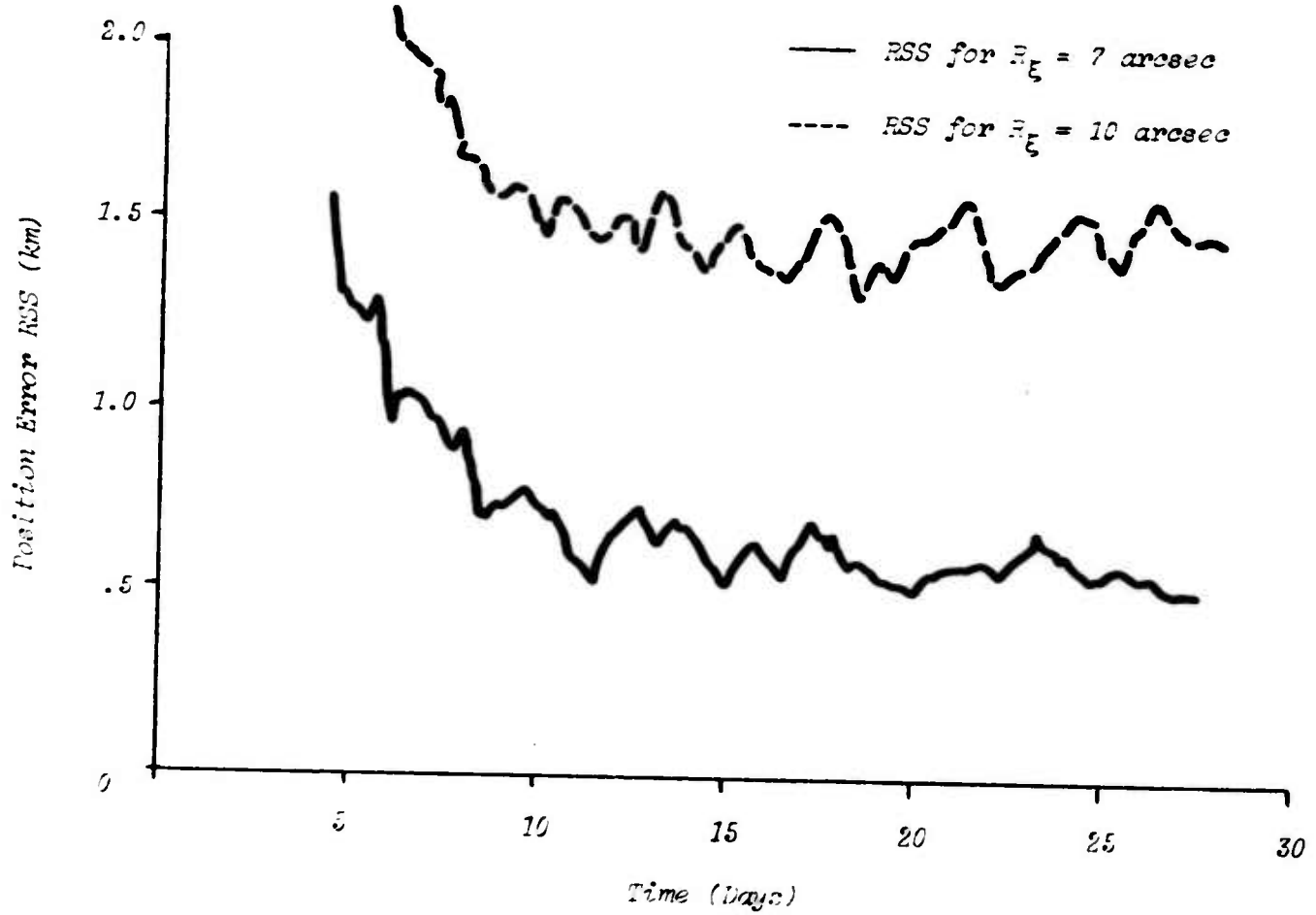


Figure 5.11 Position Error RSS with Nominal and Increased Angle Variance

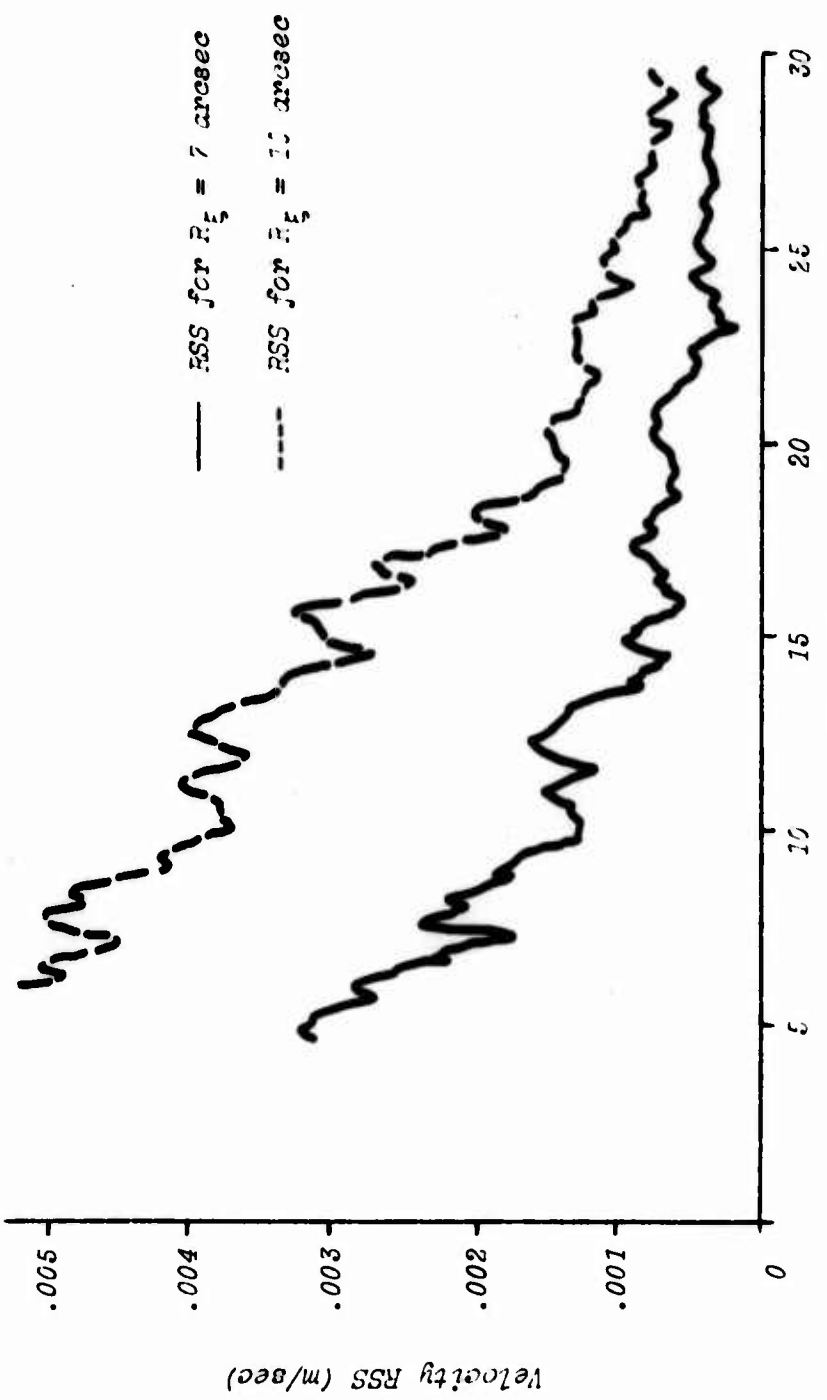


Figure 5.12 Velocity RSS for Nominal and Increased Angle Variance

equation *(1.3.21)*, letting $j = k-1$, we first make the approximation that $\Phi_{k,k-1} \simeq I$ for small Δt, so that

$$\Phi^T_{k,k-1} H^T_k R^{-1}_k H_k \Phi_{k,k-1} \simeq H^T_k R^{-1}_k H_k$$

Dropping the subscripts and employing the $\dot{\rho}$ elements of H defined in Appendix C, we have for the i - j elements corresponding to position $(i, j = 1, 2, 3)$,

$$[H^T R^{-1} H]_{ij} = [(\dot{X}_i - \dot{X}_{si}) - (X_i - X_{si})(\dot{\rho}/\rho)] \cdot [(\dot{X}_j - \dot{X}_{sj}) - (X_j - X_{sj})(\dot{\rho}/\rho)]/(\rho^2 R) \qquad (5.7.1)$$

where ρ is the range from the tracking station to the vehicle. X_i and X_{si} are elements of r and r_s respectively. For the velocity elements $(i, j = 4, 5, 6)$,

$$[H^T R^{-1} H]_{ij} = (X_i - X_{si})(X_j - X_{sj})/(\rho^2 R) \qquad (5.7.2)$$

and for the mixed terms $(i = 1, 2, 3;\ j = 4, 5, 6)$

$$[H^T R^{-1} H]_{ij} = [(\dot{X}_i - \dot{X}_{si}) - (X_i - X_{si})(\dot{\rho}/\rho)] \cdot (X_j - X_{sj})/(\rho^2 R) \qquad (5.7.3)$$

Now consider the situation where the vehicle and tracking station are in the ecliptic (heliocentric X-Y) plane, with no velocity components directed out of the plane. Clearly the information derivable from the range-rate measurement for the Z-component ($i = 3$) of position and velocity is zero. Now while this configuration is not usually the case, the values of the position and velocity Z-components of both the vehicle and the station are small enough to make the information quite small. Thus the range-rate measurement contributes only little information for the estimation of the out-of-plane components of the state. (Note that the information would be even less if station motion were not considered; this further substantiates the claim set forth in (51, 34).)

To see how the angle measurement, ξ, alleviates this problem, consider the corresponding elements of the information matrix, again for one measurement and with the approximation, $\Phi \simeq I$. For the position elements ($i, j = 1, 2, 3$),

$$[H^T R^{-1} H]_{ij} = \left[s_{xi} - \cos \xi \frac{(X_{ei} - X_i)}{|R_e - r|} \right] \cdot \left[s_{xj} - \cos \xi \frac{(X_{ej} - X_j)}{|R_e - r|} \right] / (|R_e - r|^2 \sin^2 \xi) \qquad (5.7.4)$$

Here, even when the Z-component of position is zero, the corresponding information term is, in general, non-zero. Only when

$$\xi = \cos^{-1} \left[s_z \left(\frac{Z_e - Z}{|R_e - r|} \right) \right] = \cos^{-1} \left[.3846 \left(\frac{Z_e - Z}{|R_e - r|} \right) \right]$$

is the information term zero. Now Z_e is zero, since the Earth remains in the ecliptic, and Z is very small compared to $|R_e - r|$. Hence the approximate value of $\xi = 90°$ results in a corresponding value of zero for the information term. This value can, of course, be encountered in practice. However, for the periods investigated in the simulation, this particular geometrical configuration was not encountered. (In passing we note that the values, $\xi = 0, \pi$, could lead to some numerical problems due to the division by sin ξ. However, for the navigation star selected these values of ξ cannot occur.)

For the velocity elements, the information term, $H^T R^{-1} H$, is zero since the corresponding elements of H are zero. Thus estimates of the Z-component of velocity must rely almost solely upon the information derived from the range rate. This dependence is manifested by errors in the out-of-plane velocity which are sometimes as much as an order of magnitude greater than those of the other components. However, this particular problem is partially

alleviated by accurate estimation of the position. This is true because the acceleration is, in part, a function of position, i.e., equation (6.3.1). Thus accurate estimates of r keep the nominal velocity values fairly accurate. More directly, the elements of the state transition matrix, previously assumed to be diagonal, in fact contribute a coupling of information elements to aid accurate estimates of the velocity.

While the previous discussion has been directed toward the information obtainable for the Z-components of the state, very similar conditions and explanations apply for cases where the vehicle is on or very near the X- or Y-axes. For example, initially the vehicle is on the heliocentric X-axis. Only one station, Madrid (JPL-61), is initially tracking. The vehicle is almost directly overhead, hence, $\dot{\rho} \simeq 0$, and $(Y - Y_s) \simeq 0$. Thus, only a small amount of information about $\dot{Y}$ is available since $\partial\dot{\rho}/\partial Y \simeq 0$. Similarly, for these conditions, $\dot{X} - \dot{X}_s \simeq 0$, and hence $\partial\dot{\rho}/\partial X \simeq 0$ yielding little information about X. However, information is available from the star-vehicle-Earth angle, ξ, which does aid the estimation of these otherwise locally* unobservable or nearly unobservable elements. In fact, it has been found that conditions such as these occur quite often at various times throughout the mission. The result is that without the angle measurement, ξ, extremely large estimate error values often occur.

Thus in view of the foregoing discussions, it is desirable to have an additional or supplemental measurement type. The onboard angle, ξ, appears to be a viable candidate producing useful information to aid the orbit determination process. One of its chief disadvantages is the high resolution

---

* The term local observability might be suggested for those state elements which are observable for only a subset of the total observation set. Thus a system could be described as locally observable if (1.3.11) holds where the observations $\{y_j, \ldots, y_k\}$ form a proper subset of Y_k.

demanded to produce accurate estimates. Another disadvantage is discussed later.

Model 0 - State Noise Covariance. A number of simulations were performed using various values of the diagonal state noise covariance matrix, U, (see sections 1.2 and 5.6). Typical results for three linearly related values of U over a 60-day period are shown in Figure 5.13. Here the estimation performance is reflected by the position error RSS for the following values of U:

$$U_1 = I(.16 \times 10^{-6} \quad .166 \times 10^{-5} \quad .16 \times 10^{-6})^T \quad (nm/sec^2)^2$$

$$U_2 = I(.16 \times 10^{-10} \quad .166 \times 10^{-9} \quad .16 \times 10^{-10})^T \quad (nm/sec^2)^2$$

$$U_3 = I(.16 \times 10^{-12} \quad .166 \times 10^{-11} \quad .16 \times 10^{-12})^T \quad (nm/sec^2)^2$$

Examining the figure, as the state noise covariance is decreased, we begin to see divergence of the state estimate. In fact, for all cases, including that corresponding to the largest value, U_1, divergence is at least starting near the end of the simulation period.

A particularly interesting feature is the apparent trade-off in the values of U and the maximum estimation accuracy obtained. This is shown by the large dips in the RSS curves at about 10 days. For the smallest value, U_3, the RSS curve has the greatest dip; for U_1 the dip is smallest. An intuitive explanation of this phenomenon is based on the following reasoning. Initial filter operation produces estimates of greater and greater accuracy as more observations are taken. If the value of U is small, the error covariance decreases rapidly, producing more nearly optimal estimates and thus, for a short time, more accurate estimates. However, eventually the covariance becomes so small that it is unable to cope with the error buildup, and divergence occurs. On the other hand, for a larger U the error covariance does not decrease as rapidly. Hence, the filter is not operating near the optimum,

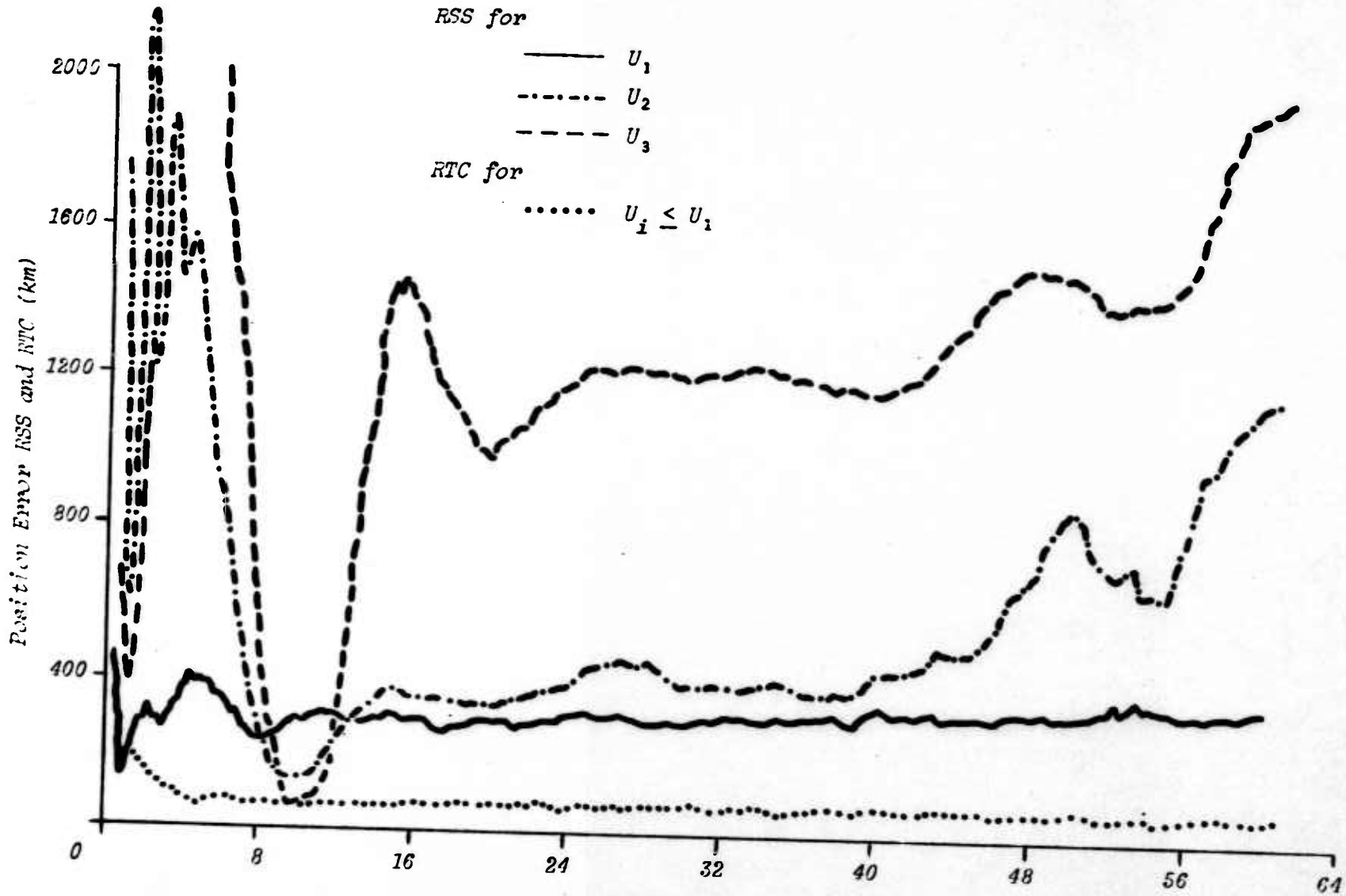


Figure 5.13 Position Error, Model 0

and estimate accuracy is therefore not as good -- at least during the short period around 10 days. However, the presence of a larger U keeps the filter operating longer before divergence occurs.

Next note the dotted curve which corresponds to the maximum RTC, i.e., that for U_1. Its average value is approximately 70-80 km, a value which does not reflect the error RSS of 300-400 km. (Although they are not shown, the RTC's corresponding to U_2 and U_3 all lie below that for U_1.) This suggests that the value of U be increased.

Such an increase produces the estimation results shown in Figures 5.14 and 5.15. The increased elements of U have the values,

$$q_x = q_z = .16 \times 10^{-4} \ (mn/sec^2)^2$$

$$q_y = .166 \times 10^{-3} \ (mn/sec)^2$$

These figures show position and velocity RSS and RTC curves for both the nominal acceleration error and an increased error, to be discussed later. In Figure 5.14 the position error RSS is seen to be greater than that obtained in Figure 5.13. However, the RTC curve is a better measure of the accuracy, even though it does not bound the RSS curve. Further, the estimate remains fairly stable throughout the simulation period of 60 days. This is further supported by examining Figure 5.15. Here the solid curve represents the velocity error RSS over the 60-day simulation period. Although there are some fairly large error peaks during the latter 30 days, overall the velocity error RSS appears generally to be free of divergence during this period.

Figures 5.14 and 5.15 also show the estimation performance for the case where the thrust acceleration error is increased by a certain amount at 30 days into the mission. This increase is produced by changing to the following acceleration error simulation values:

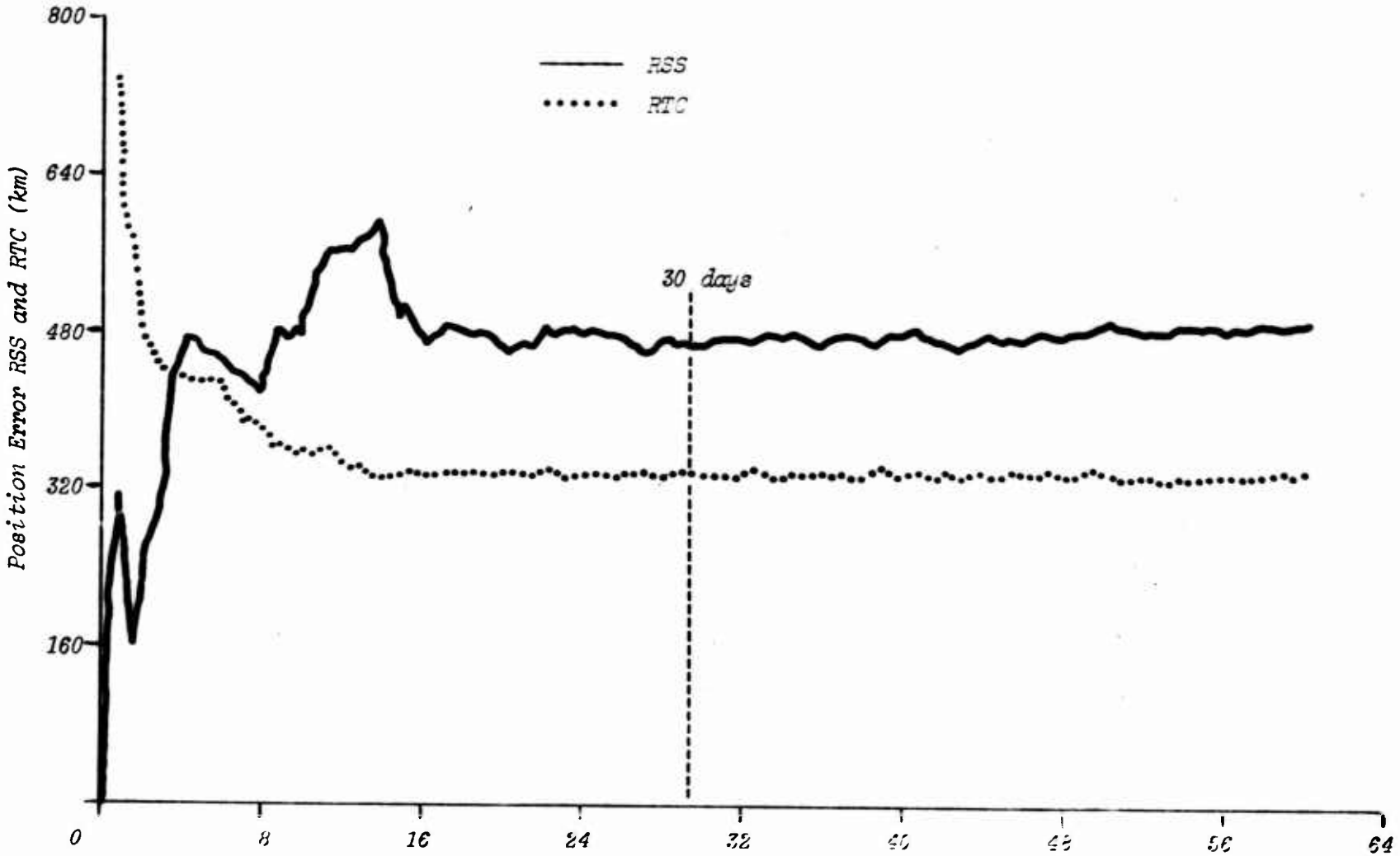


Figure 5.14 Position Error, Model O, Nominal and Increased Acceleration Error at 30 Days

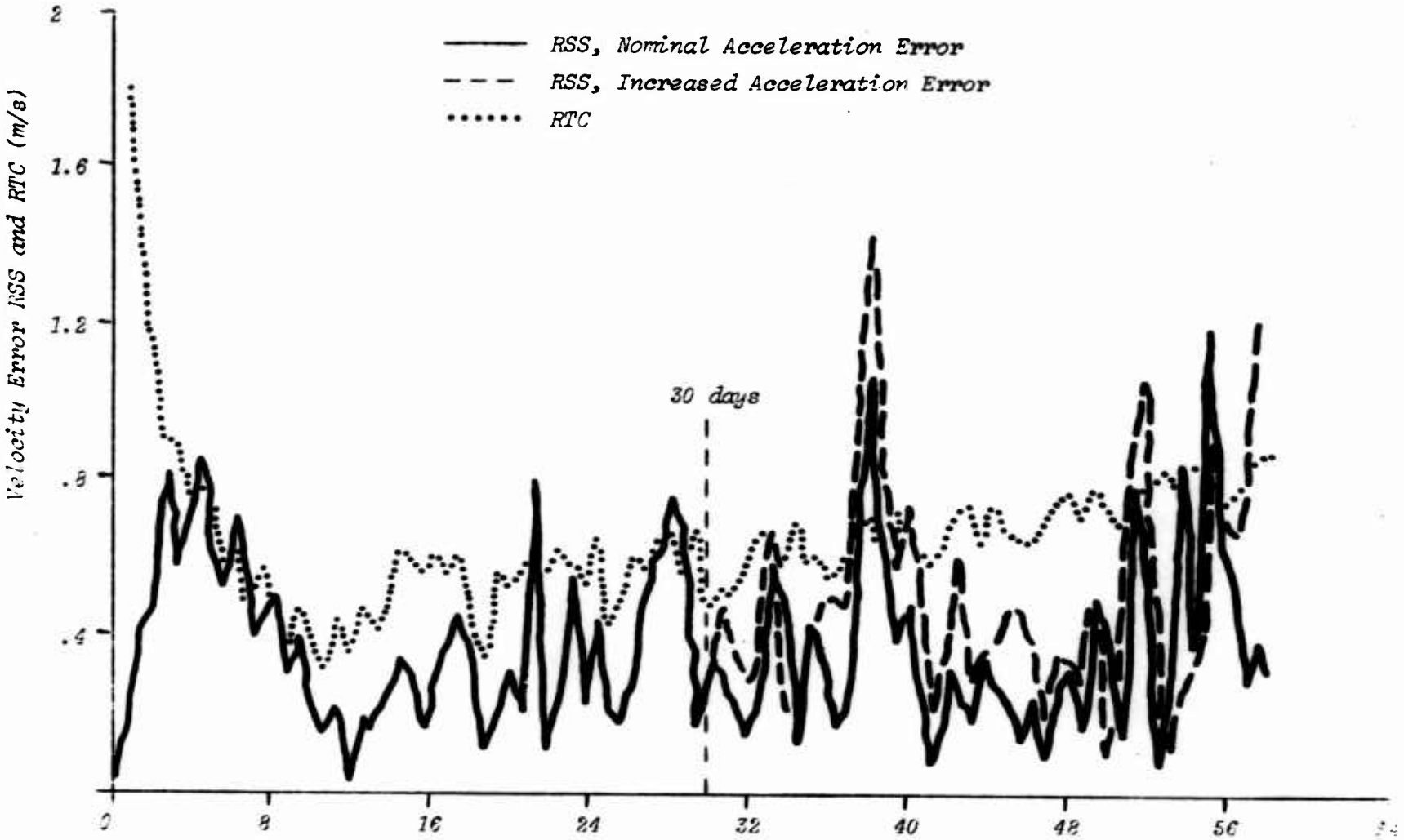


Figure 5.15 Velocity Error, Model O, Nominal and Increased Acceleration Error at 30 Days

$\bar{\gamma} = .026175\ rad$

$\delta a_0 = .0027\ m/sec^2$

This corresponds to a 50% increase in the nominal values. A plot of the corresponding y-component of thrust acceleration error is shown in Figure 5.16. Figures 5.14 and 5.15 thus reflect the ability of the algorithm to handle changes in the thrust acceleration error. For the corresponding position error RSS, the differences between the nominal and increased thrust acceleration error are essentially nonexistent, and are therefore shown together as the solid curve. The dotted curve is the RTC for both also. In Figure 5.15, however, the velocity error RSS curves are slightly different, with the RSS for the increased acceleration being given by the dashed curve. As expected, the error is slightly larger than that for the nominal acceleration error simulation. As in Figure 5.14, the dotted curve represents the velocity error RTC for both cases. However, a curious aspect is the behavior of the RTC curve. This is seen as a generally increasing trend from about 10 days. The data also reflect a similar, but considerably less detectable, behavior for the position RTC. One possible explanation for this is the fact that as the vehicle moves farther from the Earth, the angle measurement, ξ, becomes less sensitive to changes in position. For example, initially the spacecraft is approximately 9.25×10^5 km from the Earth. The standard deviation of 7 arcsec for the onboard angle measurement corresponds to an arc length at this distance of approximately 31.4 km. By 60 days into the mission, the spacecraft is approximately 1.53×10^7 km from the Earth: the corresponding arc length is approximately 519 km. While this behavior could be seen more rigorously in terms of the information matrix, the effect is clear: less information is available from the onboard measurement angle as the distance from the Earth is increased.

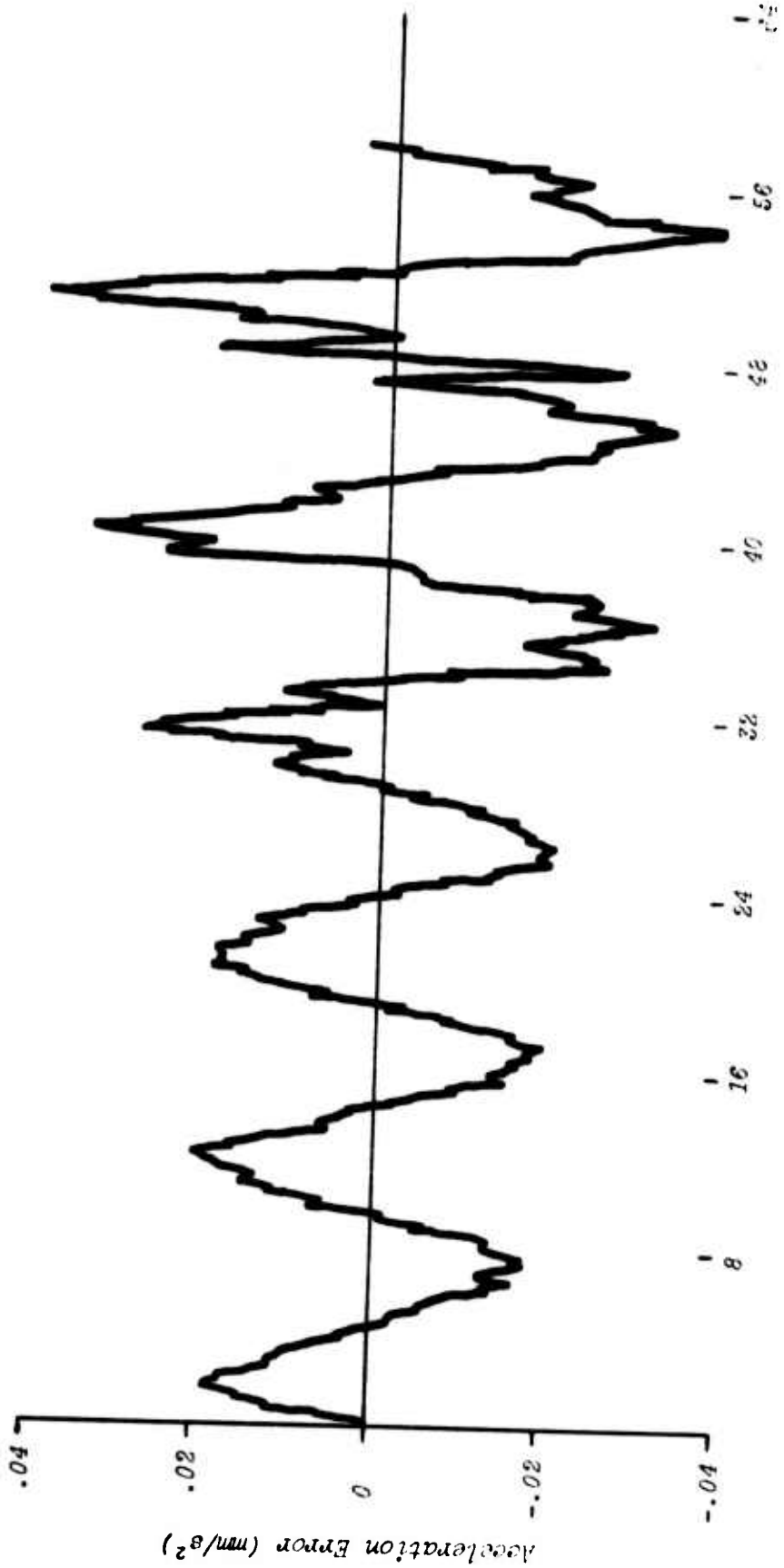


Figure 5.16 Acceleration Error (y-axis), Increase at 30 Days

Models 1 and 2. Figures 5.17, 5.18, and 5.19 show the estimation performance for each of the other four approximating models. For these cases, the thrust acceleration error parameters are increased to the same values as for Model 0, above. Further, each case employed values of q_l as indicated below.

$$\text{Model 1a} \begin{cases} q_x = q_z = .16 \times 10^{-4} \ (mm/sec^2)^2 \\ q_e = .16 \times 10^{-10} \ (mm/sec^3)^2 \end{cases}$$

$$\text{Model 1b} \begin{cases} q_x = q_z = .16 \times 10^{-4} \ (mm/sec^2)^2 \\ q_e = .5 \times 10^{-14} \ (mm/sec^3)^2 \\ q_\alpha = .q \times 10^{-35} \ sec^{-4} \end{cases}$$

$$\text{Model 2a} \begin{cases} q_x = q_z = .16 \times 10^{-4} \ (mm/sec^2)^2 \\ q_g = .3 \times 10^{-24} \ (mm/sec^4)^2 \end{cases}$$

$$\text{Model 2b} \begin{cases} q_x = q_z = .16 \times 10^{-4} \ (mm/sec^2)^2 \\ q_g = 10^{-22} \ (mm/sec^4)^2 \\ q_\beta = .5 \times 10^{-30} \ sec^{-6} \end{cases}$$

Figure 5.17 shows the position error RSS for each of the model configurations. As expected, the estimation accuracy generally is improved over that for Model 0. Most remarkable is the accuracy with which Model 2b yields performance estimates. This is represented by the solid curve, and corresponds to an average position RSS of about 60 km (time ~ 10 days). This approaches a factor of 5 reduction in the RSS values for the other models. It is important to remember, however, that Model 2b corresponds to the correct structure of the actual thrust acceleration error magnitude, and thus one would expect superior performance. We remark that Model 2b is not quite

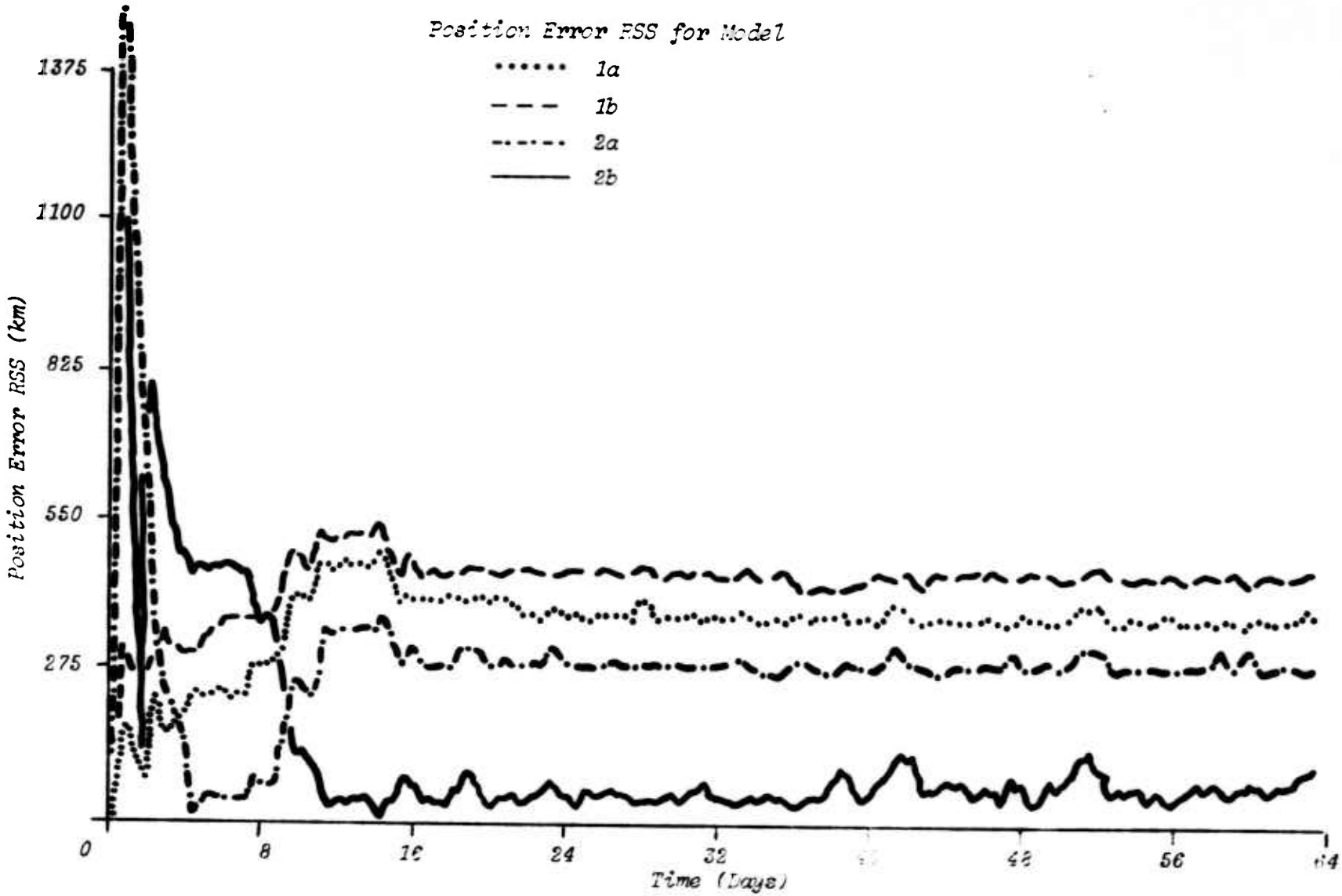


Figure 5.17 Position Error, Models 1 and 2, Increased Acceleration at 30 Days

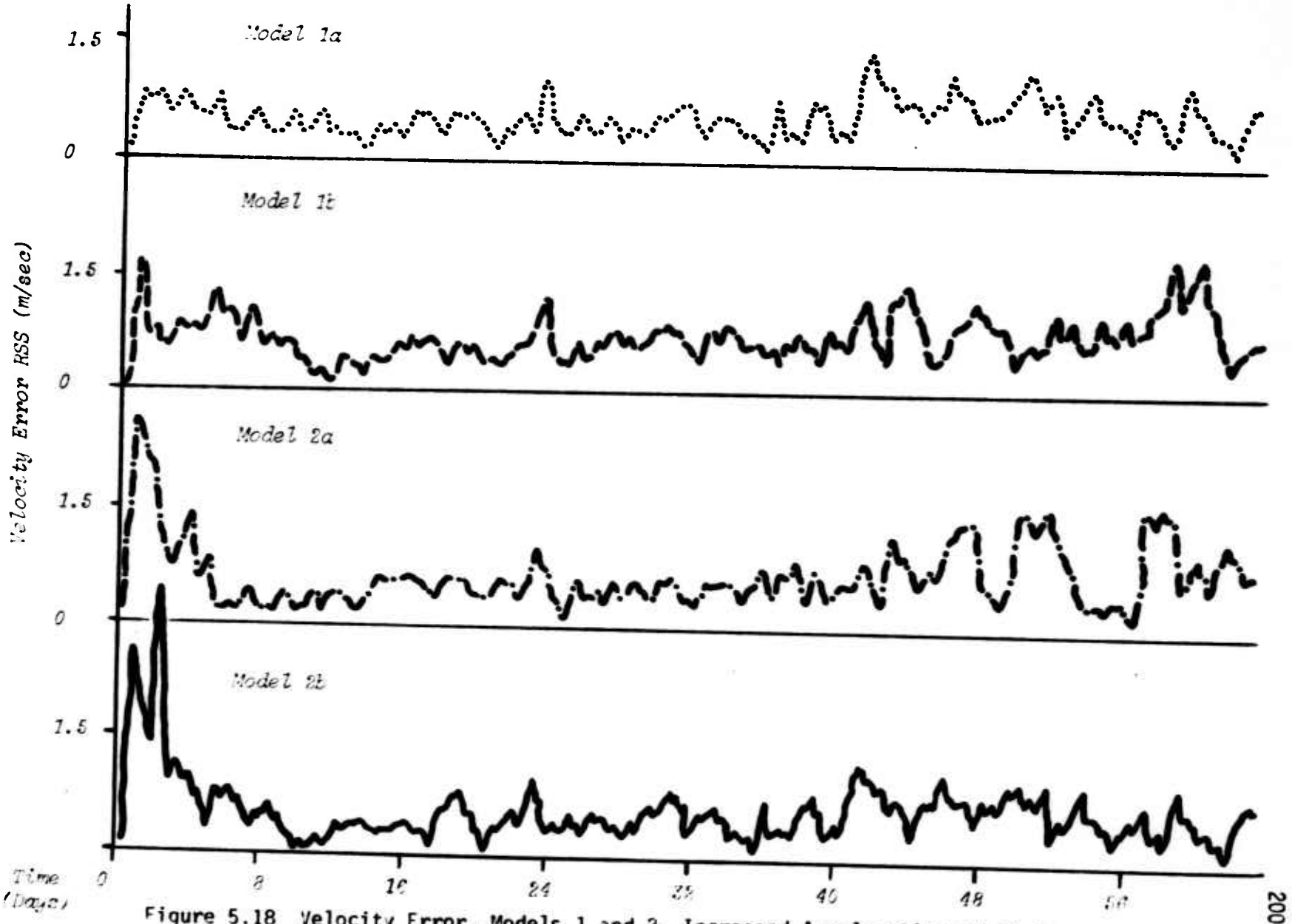


Figure 5.18 Velocity Error, Models 1 and 2, Increased Acceleration at 30 days

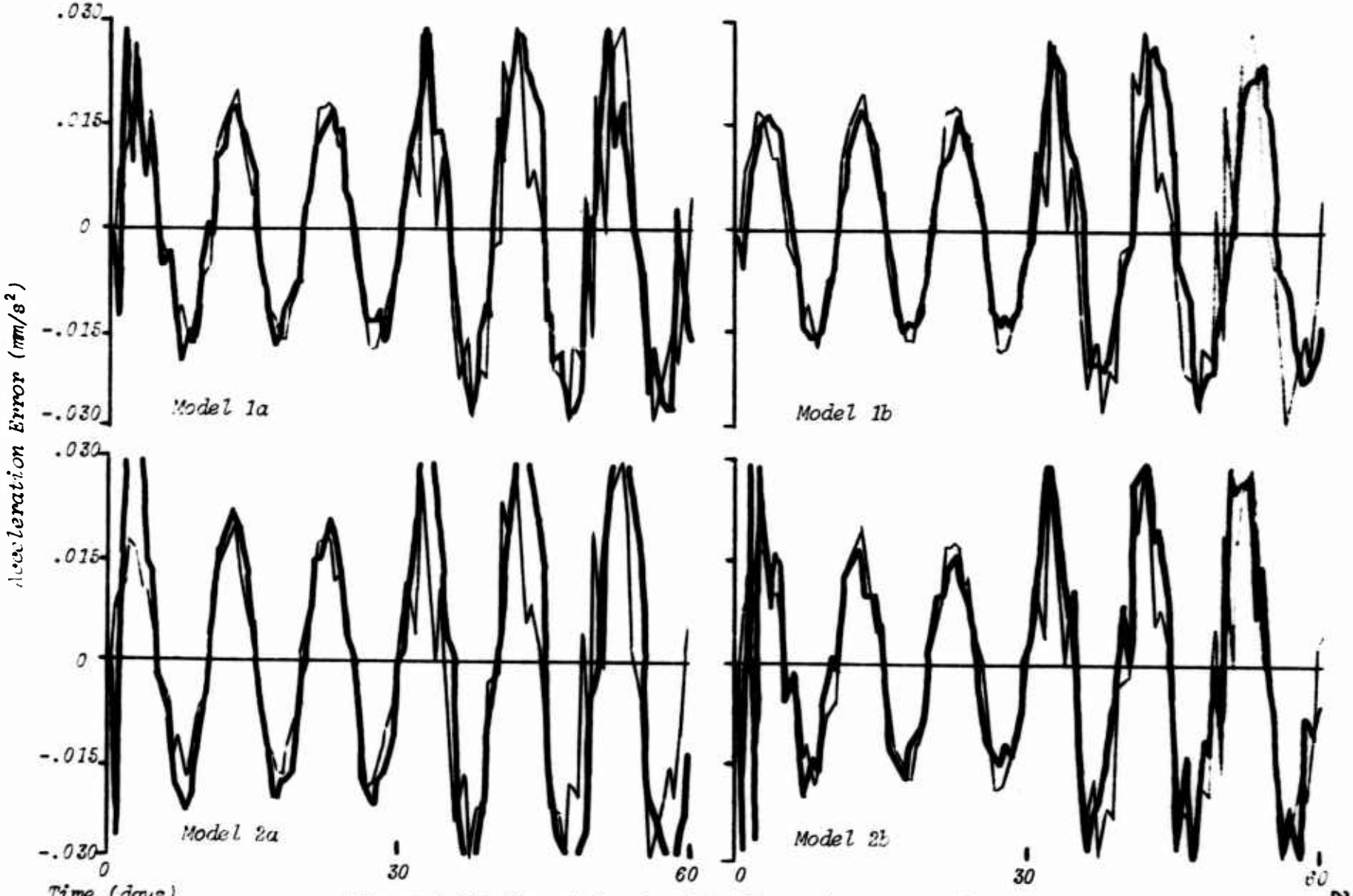


Figure 5.19 Thrust Acceleration Error (y component), Increased Error at 30 Days

the exact structure for the y-component of acceleration error, however. This is due, of course, to the fact that the thrust acceleration error vector deviates from its nominal orientation so that Model 2b is attempting to approximate the y-axis projection of the thrust acceleration. Further, we have employed a differential form, and consequently truncation errors in the numerical integration algorithm contribute what appears to the estimation equations to be a form of model error. It is these aspects which require the use of a small state noise covariance matrix (also see (11)).

The next most accurate performance is that provided by Model 2a, simple linear form. As shown by the alternately dot-dashed curve, the average position RSS is approximately 300 km (for time > 10 days).

The two first order models, 1a and 1b, produced acceptable estimation accuracy, but not as good as expected. In fact, Model 1b produced a position RSS curve of about 450 km, some 80 km greater than that for Model 1a. One possible reason for this is that too small a value for one or a number of the state noise covariance elements was used. This suggestion is supported by the fact that the associated error RTC data indicated an approximate steady state value of 280 km, a value somewhat below the actual RSS curve. In all cases a considerable number of runs were made for each model to obtain appropriate values for the state noise covariance matrix elements. This process is tedious, time-consuming, and expensive, and for this reason, it was decided, following some effort, not to pursue this "tuning" process further. Clearly, one would expect that more accurate results could be obtained by varying the q_l values until the position error RTC and RSS become approximately the same.

For each of the other models, the indicated values of q_l resulted in RTC's which coincided with or exceeded their respective RSS data. The model producing an error RTC greater than its RSS data was Model 2b. In particular

the RTC data indicated an average value of approximately 295 km. Presumably further tuning of the filter would result in even greater accuracy for Model 2b. For the reasons cited above, this was not done.

Figure 5.18 shows the velocity error RSS curves for each of the models. Here the performance appears to be nearly the same for each. However, these curves are in general more erratic than in Figure 5.19. In particular, toward the end of the simulation interval the effects of the increased thrust acceleration error are rather pronounced. In spite of this, there does not seem to be any indication of a tendency toward divergence; hence, for all these models, the estimates are stable over the entire interval. Further, in all cases the corresponding RTC data bounded or coincided with the RSS data.

Figure 5.19 shows the estimates of the thrust acceleration y-component for each model. These estimates are those which resulted in the estimation performances of Figures 5.17 and 5.18. Each estimated acceleration curve represented by the heavy, solid line is identified with the corresponding model number. The estimated curves are superimposed over a lighter curve, the true y-component of thrust acceleration error. As seen from the figure, all four models perform admirably in representing this component of acceleration error. Particularly notable is the fact that all methods adapt to the increase in the acceleration error at 30 days. Now, the plots for Models 1a and 2b do not appear to be very different (although the position and velocity RSS's in Figures 5.17 and 5.18 indicate otherwise). On the other hand consider the respective curves for Models 1b and 2a. Close examination of the Model 1b curve reveals that a number of the actual error peaks are not matched by the approximating curve. This is possibly due to a smaller than necessary state noise covariance matrix and presumably the estimates would be improved if this matrix were increased in value. This further supports

the suggested explanation why Model 1b yields the largest position error RSS. Next, examining the Model 2a curve reveals that the estimated values near the error peaks are larger than the actual error (the breaks in the approximating curve indicate where the values extend outside the limits of the graph). In view of the previous suggestion concerning Model 1b that the state noise covariance is too small, the implication here is that the corresponding covariance is too large. In this case, however, it has been found that the position error RTC data coincides very closely with the RSS curve. Further, the fact that the Model 2b simulation yielded RTC values which were well above its RSS values, tends to discount this explanation for over-estimating the peaks. The actual reason for this behavior is not clear. Two possible explanations are offered. First, it is possible that generally over-estimating the acceleration error simply results in orbit determination performance which is more accurate than that resulting from under-estimating the error. Second, the plots shown are generated using approximately every fourteenth or fifteenth point. Thus many values occur which are not shown, and *on the average*, it is possible that considerably better estimates of the error are obtained for Model 2a than for Model 1b.

In view of the foregoing discussions, it appears that generally the second order models, 2a and 2b, are superior in representing thrust acceleration errors of the type encountered in the simulations. Further, the importance of exact modeling to the maximum extent practical is clearly illustrated by the exceptional performance of Model 2b.

Sequential Estimation of the State Noise Covariance. We have seen rather accurate performance using an *a priori* constant state noise covariance matrix associated with each of the models. However, in view of the many numerical simulations which invariably must be performed in order to "tune"

the state (vector) filter, a natural alternative is to employ the sequential state noise covariance or Q-filter to aid in this process. This has been quite successful as we now proceed to show.

First, some comments dealing with general performance are in order. Initially, numerous runs were made to investigate the effectiveness of the sequential Q-filter. In these cases operation of both the ordinary state filter and the Q-filter was simultaneously initiated. Results were obtained for all models and were less than satisfactory. In all of these cases, the performance was worse than that obtained by simple tuning of the state filter using *a priori* constant values of Q. Attempts were made to improve the performance of the Q-filter by employing different *a priori* values of both Q and the Q-error covariance, S; using different constant values of the observation residual error variance, T; and using the "sliding window" modification to the T-estimator, equation *(3.6.36)*. These attempts met with little success. After exhausting such approaches, a re-examination of the strategy for employing both the state and Q-filters was made. This revealed that possibly the initial transient operation of the state filter was having a detrimental effect upon the Q-filter performance. To investigate this, runs were made in which in which initiation of the Q-filter operation was postponed for the transient period of state filter operation, approximately 5 days. The results were highly encouraging.

To illustrate this strategy, cases for Models 1b and 2b were run. Recall from previous discussion that the results obtained in Figures 5.17 - 5.19 suggested that the state noise covariance values appeared to be too small and too large, respectively, for these models. For this reason, it was felt that the Q-filter could improve the values of Q, and hence the estimation performance. For both models, simulations were run for 30 days using the

initial data given in Table 5.1. In addition, *a priori* values of Q were taken to be those of U corresponding to the results obtained in Figures 5.17 - 5.19* (see page 198). The initial values of the Q error covariance matrix, S, were selected to make S_0 diagonal. The S_0 elements corresponding to diagonal elements of Q were chosen approximately as the square of these *a priori* Q values. The S_0 elements corresponding to off-diagonal or correlation terms of Q were taken (approximately) as the product of the *a priori* diagonal Q elements corresponding to the row and column of the correlation term. Thus, at $t = 0$,

$$s_{ii} \simeq q_{jj}^2 \ , \quad s_{kk} \simeq q_{mm}\, q_{nn} \qquad (5.7.5)$$

Taking this approach was an arbitrary choice. However, it was found that relatively accurate values for S_0 are necessary, and the rule-of-thumb given by *(5.6.5)* was found to be a viable approach. In view of this, the *a priori* S values are given in Table 5.2 for each of the two b models.

Because of the extremely small values of many of the s_{ii}, different units are used in carrying out these simulations, namely, millions of meters and millions of seconds. This is done in order to avoid numerical difficulties. In order to test the adaptability of the Q filter, increased acceleration errors were simulated at 15 days. This was done by again increasing $\bar{\gamma}$ and δa_0 by 50% to

$$\bar{\gamma} = .026175 \ rad$$

$$\delta a_0 = .0027 \ m/sec^2$$

In determining the observation residual error variance, r, the sliding window modification, equation *(3.6.36)* was used with the small *a priori* value of

---

* Recall the discussion in section 5.5 associated with the computation of Γ. Because Γ itself is required, equation *(1.2.30)* is used, and hence we assume, approximately, $Q = U$.

Reproduced from best available copy.

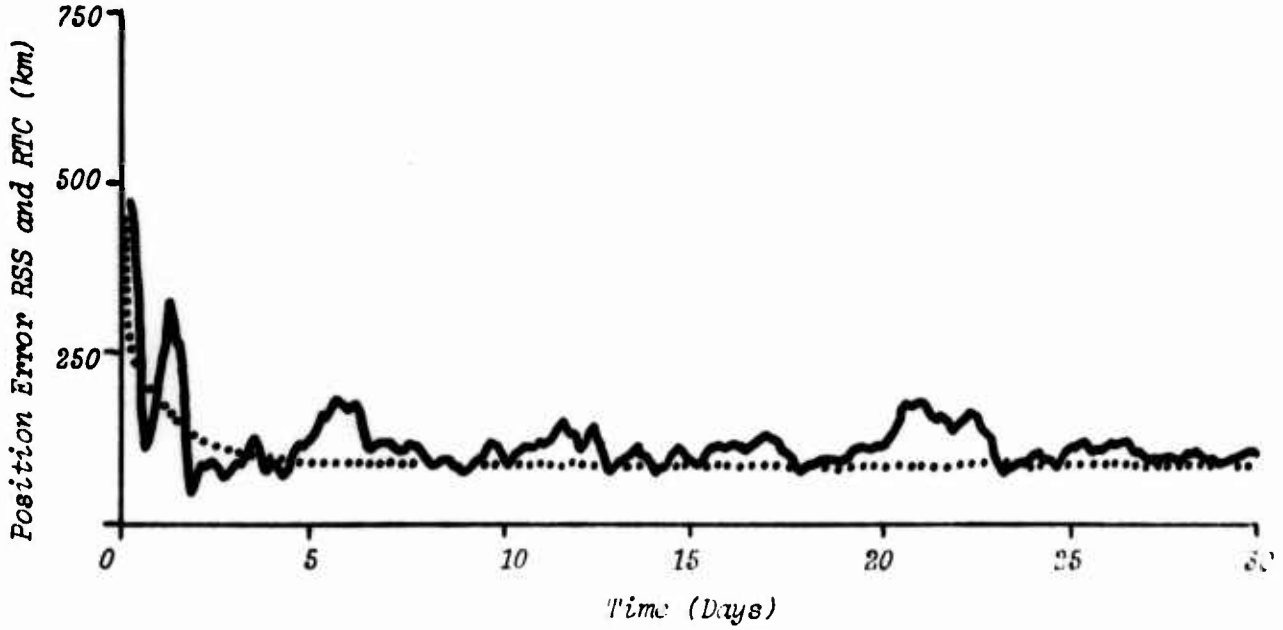
Diagonal S Element	Corresponding Q Element	Variance Values for Model 1b	Model 2b
s_{11}	$q_{11} = q_x$	10^{-10} $(mm/sec^2)^4$	10^{-10} $(mm/sec^2)^4$
s_{22}	q_{12}	10^{-10} $(mm/sec^2)^4$	10^{-10} $(mm/sec^2)^4$
s_{33}	q_{13}	10^{-19} $(mm^2/sec^5)^2$	10^{-27} $(mm/sec^3)^4$
s_{44}	q_{14}	10^{-40} $(mm/sec^4)^2$	10^{-35} $(mm/sec^5)^2$
s_{55}	$q_{22} = q_z$	10^{-10} $(mm/sec^2)^4$	10^{-10} $(mm/sec^2)^4$
s_{66}	q_{23}	10^{-19} $(mm^2/sec^5)^2$	10^{-27} $(mm/sec^3)^4$
s_{77}	q_{24}	10^{-40} $(mm/sec^4)^2$	10^{-35} $(mm/sec^5)^2$
s_{88}	$q_{33} = q_e,\ q_g$	10^{-28} $(mm/sec^3)^4$	10^{-44} $(mm/sec^4)^4$
s_{99}	q_{34}	10^{-49} $(mm/sec^5)^2$	10^{-52} $(mm/sec^7)^2$
$s_{10\ 10}$	$q_{44} = q_\alpha,\ q_\beta$	10^{-70} sec^{-8}	10^{-60} sec^{-12}

Table 5.2 A *Priori* S Values

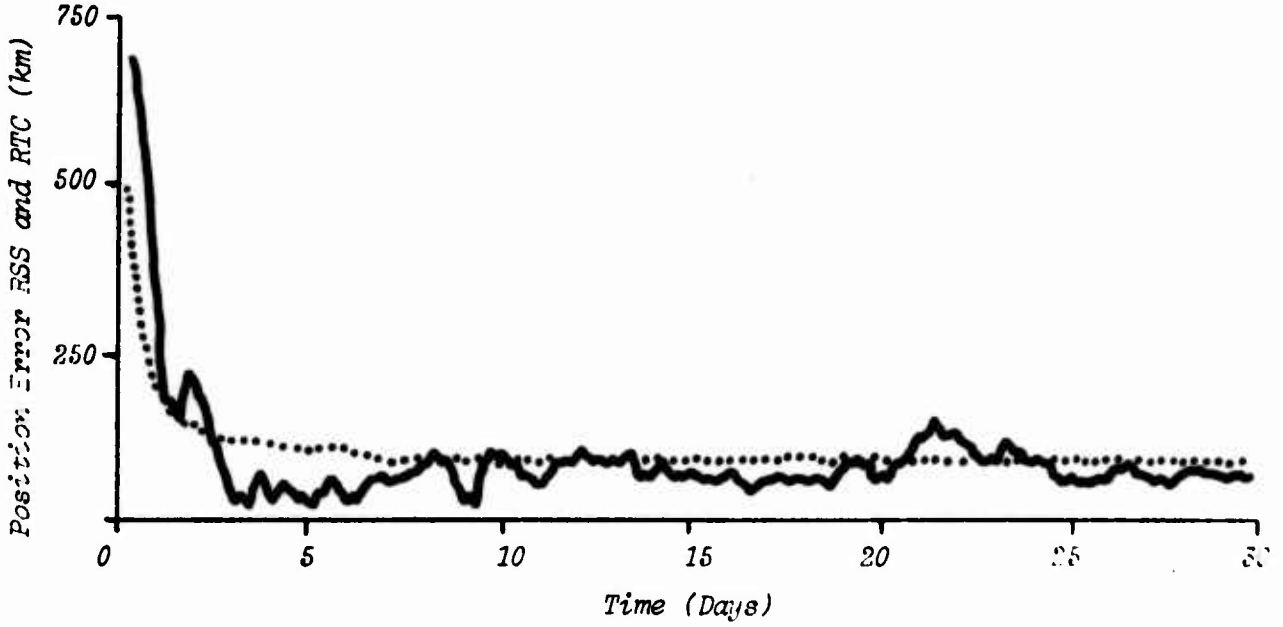
$T_0 = 10^{-14}$ (nm/sec)^4. Twenty residuals were used as the residual batch size, i.e., $N = 20$, in (3.6.36). Finally, in carrying out the simulations, the Q filter was not operated until the fifth day of mission time in order to avoid the transient effects of state filter operation.

Figure 5.20 shows the position RSS and RTC for both b models: 5.20a for Model 1b, and 5.20b for Model 2b. In both cases the performance is quite good. Figure 5.20a shows a particularly significant increase in performance over that shown in Figure 5.17. The RSS curve has been lowered to an approximate mean of 125 km, a considerable improvement over the previous 450 km of Figure 5.17. Further, the RTC curve is more representative of the actual RSS error, although it appears to be slightly low. For Model 2b there does not seem to be any noticeable improvement in the estimate as reflected in the RSS curve. Both here and in Figure 5.17 the RSS has a mean of around 60 km. However, the corresponding position error RTC is now an accurate representation of the position error, and adequately bounds the actual error RSS. Note that for both models the filters easily handle the 50% acceleration error increase at 15 days.

Results obtained for the velocities are shown in Figures 5.21a and 5.21b. The performance is comparable to that shown in Figure 5.18. However, for Model 1b the velocity error RSS curve has some peaks which are larger than those of Figure 5.18b. These peaks seem to be decreasing in amplitude after the largest one at 21 days, and therefore do not suggest divergence. However, the RTC is somewhat lower than the RSS curve, and it is possible that divergence may appear at some later time. As before, the results for Model 2b are particularly good, although here also there is a large error peak at about 21 days. Again the RSS decreases, and, except for the peak, the RTC mean of about .2 m/sec adequately reflects the true error.

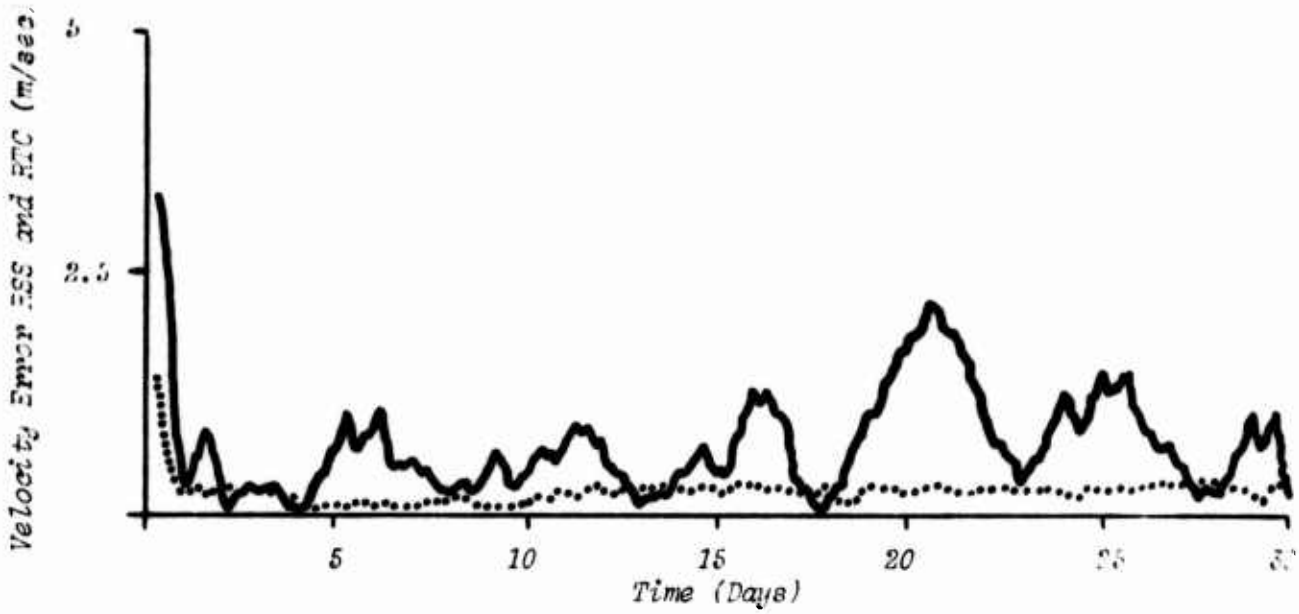


5.20a Position Error RSS and RTC for Model 1b

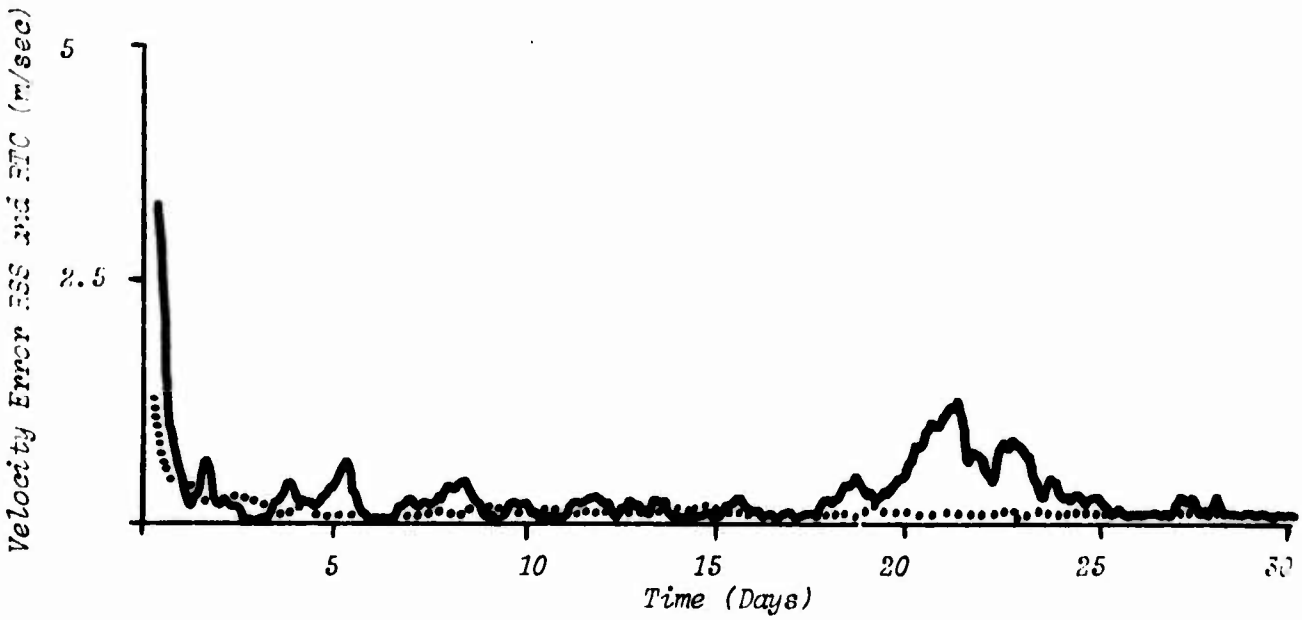


5.20b Position Error RSS and RTC for Model 2b

Figure 5.20 Position Error RSS and RTC With Sequential Q Estimation



5.21a Velocity Error RSS and RTC for Model 1b

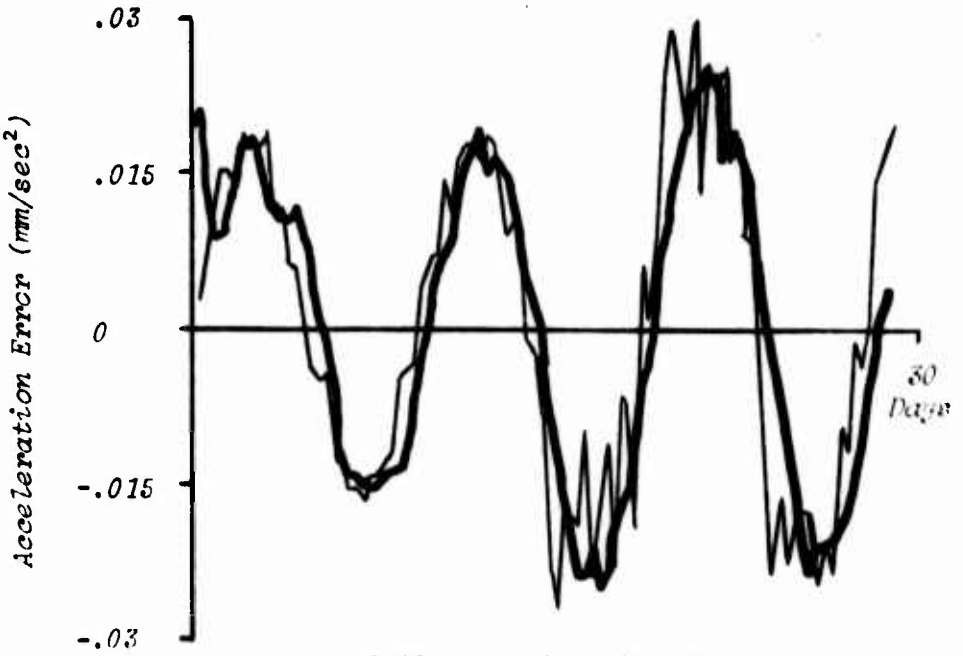


5.21b Velocity Error RSS and RTC for Model 2b

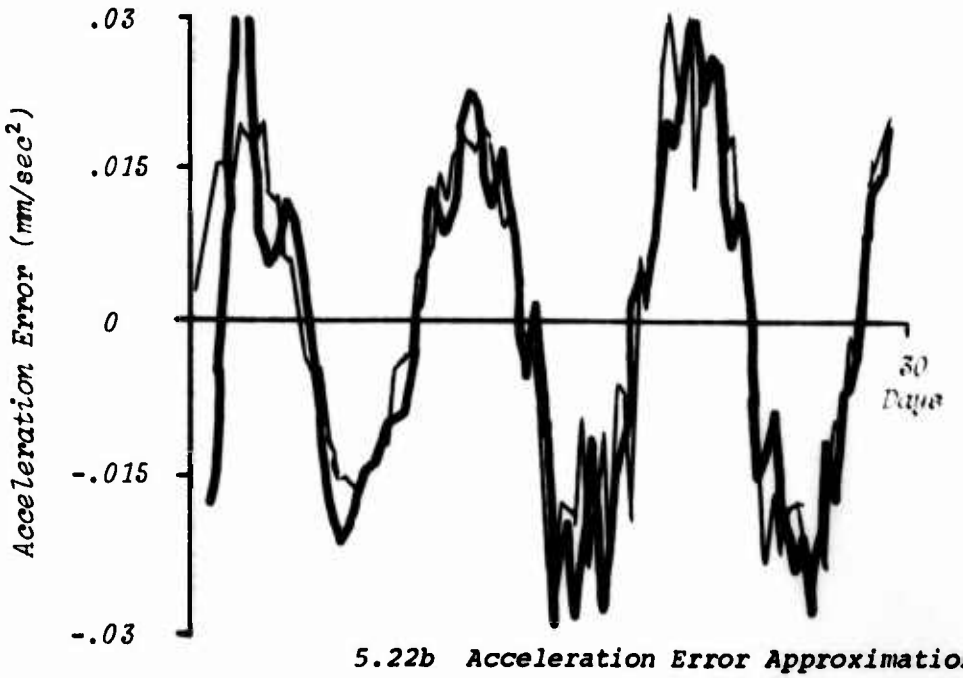
Figure 5.21 Velocity Error RSS and RTC
With Sequential Q Estimation

It is entirely possible that the peaks in the error RSS curves may be due to operation of the filters as they attempt to adapt to the acceleration error. Figures 5.22a and 5.22b suggest this possibility as they show the true and approximated thrust error accelerations. In particular, Figure 5.22a shows that the Model 1b error approximation is slightly out of phase with the true error. Further, it is smoother and does not reflect the more erratic behavior of the actual thrust acceleration error. On the other hand, Model 2b does a reasonably good job of approximating the true error. In either case there are rapid alternations at the true error maxima and minima which are difficult to approximate and which are therefore quite capable of resulting in the peaks in the position and velocity error RSS'.

Figures 5.23 and 5.24 are plots of appropriate RTC values for the state noise covariances. These figures illustrate quite well the expected changes in the state noise covariance values. Figures 5.23a and 5.23b show the respective Model 1b and 2a state noise covariance RTC curves for the orbital frame x-z components of error. Note that both curves reflect the constant *a priori* values. At 5 days there is a rapid *increase* in the RTC value for Model 1b and a *decrease* in the RTC for Model 2b. This behavior is exactly that expected to improve the position and velocity RSS'. Also notable is the increase in these Q estimates at about 15 days, corresponding to the increased thrust acceleration error. Thus the Q filter is actually adapting its estimates to account for these error increases. This behavior can also be seen in Figures 5.24. Figure 5.24a shows the state noise covariance RTC, $\sqrt{\hat{q}_e}$, corresponding to the noise term in the approximating differential equation *(5.5.5)*. Likewise, Figure 5.24b shows the covariance RTC, $\sqrt{\hat{q}_g}$, corresponding to the noise term in *(5.5.10)*. We again see the constant *a priori* Q values for both cases during the first 5 days of the mission. At

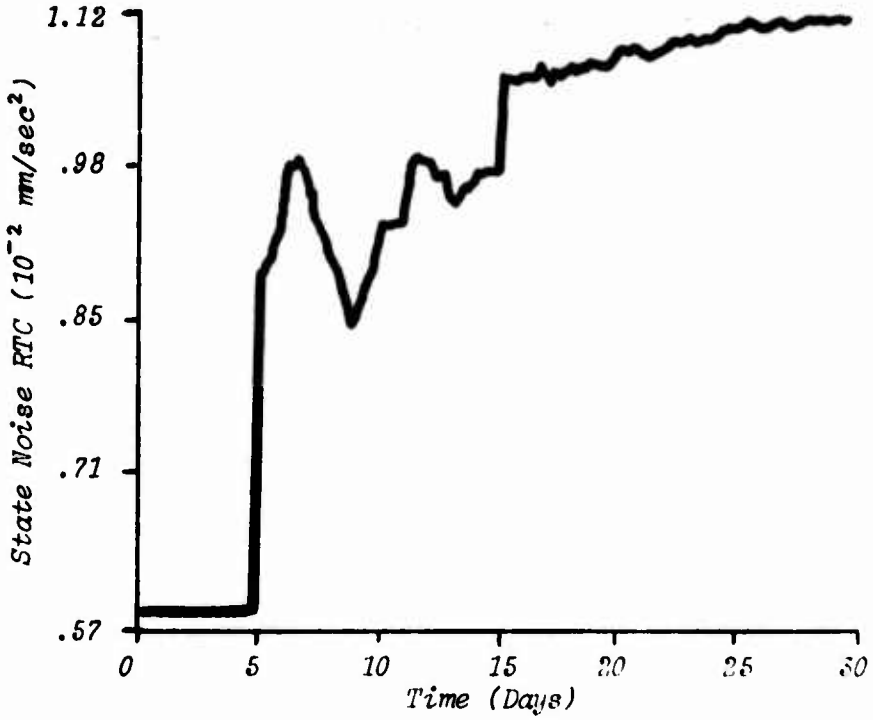


5.22a Acceleration Error Approximation by Model 1b

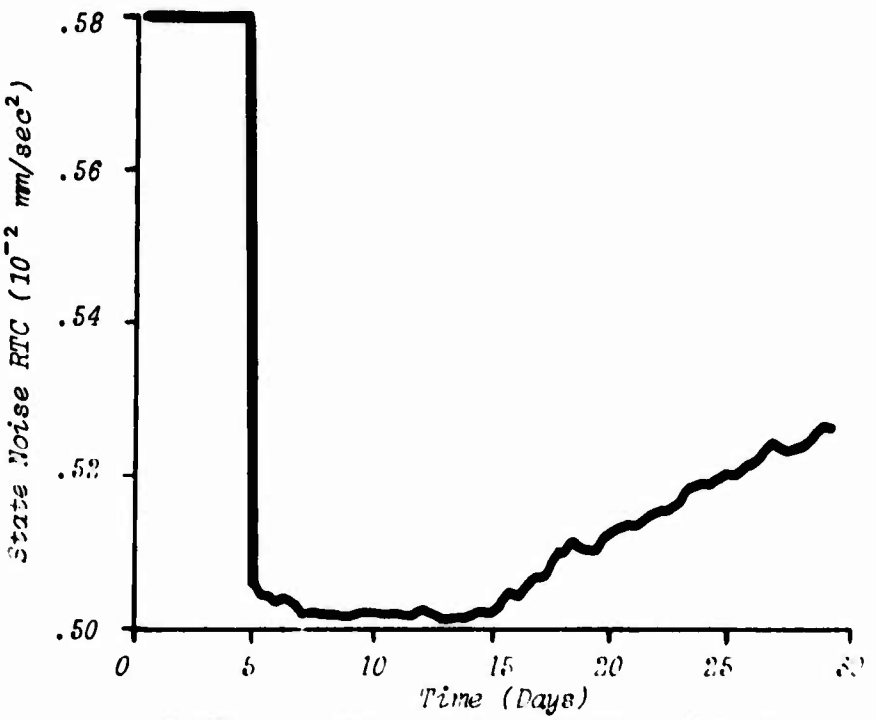


5.22b Acceleration Error Approximation by Model 2b

Figure 5.22 Acceleration Error Approximations Employing Sequential Q Estimator

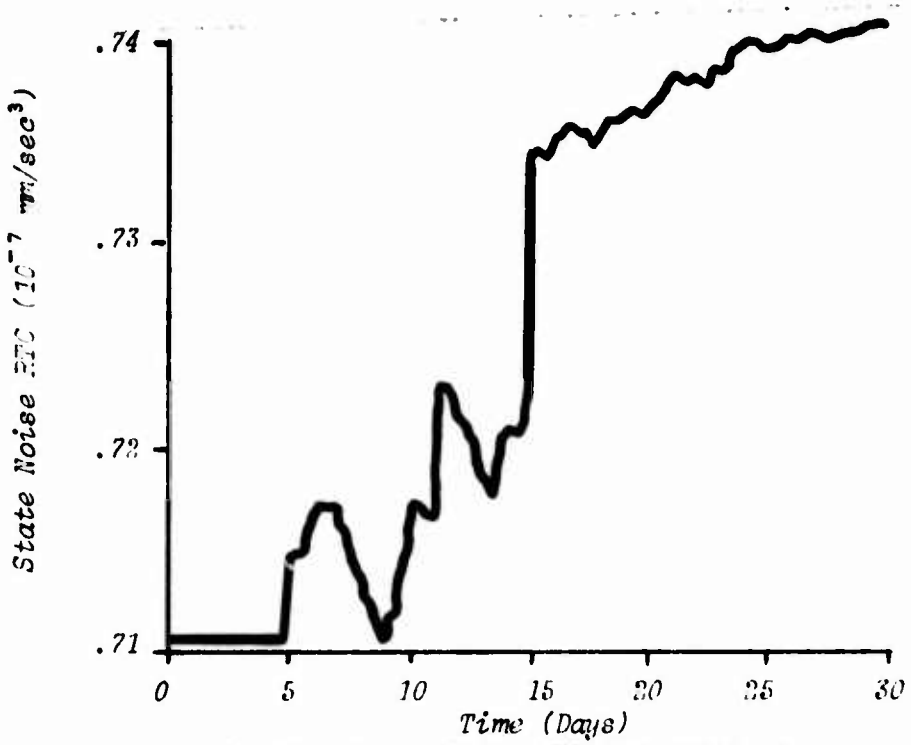


5.23a *x-z State Noise RTC for Model 1b*

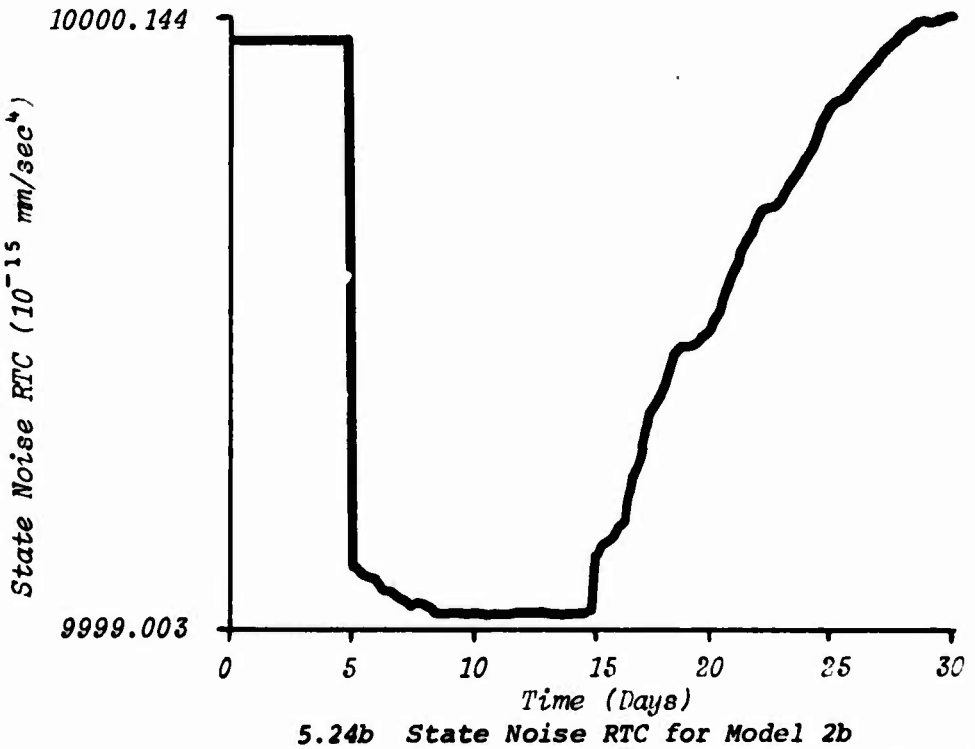


5.23b *x-z State Noise RTC for Model 2b*

Figure 5.23 State Noise RTC of Orbital Frame x-z Components



5.24a State Noise RTC for Model 1b
Acceleration Error Derivative, $\dot{a}$



5.24b State Noise RTC for Model 2b
Acceleration Error 2nd Derivative, $\ddot{g}$

Figure 5.24 State Noise RTC for Derivatives in Models 1b and 2b

this point changes occur with $\sqrt{\hat{q}_e}$ increasing and $\sqrt{\hat{q}_g}$ decreasing as expected. Again at 15 days both estimates begin growing as the Q filter attempts to account for the increased model error.

For the model parameters, α and β, the corresponding state noise covariance estimates retained their *a priori* values with no discernible changes as the estimation proceeded. This is attributed to the very small *a priori* values of $s_{10\ 10}$ corresponding to these parameters (Table 5.2). Recall from Chapter 4 that too large a value of Q for the model parameter, c, resulted in c's following the observations, thus destroying the correlation effects. In view of this, the suggestion is made that it is better, within parexic bounds, to let such parameters tend more to their constant values. This strategy is further motivated by the initial results obtained when both the state and Q estimations were initiated simultaneously. Thus initial values of S normally should be chosen as nearly representative of the true uncertainties in Q as possible. An interesting aspect of the Q filter performance was noted in using constant values of the residual variance, T, as opposed to estimating them via *(3.6.36)*. Results indicated that variations in T by as much as three orders of magnitude produced essentially no overall change in estimation accuracy from that obtained in the foregoing cases. Generally, however, it is recommended that T be estimated, thus eliminating additional *a priori* guesswork.

The main disadvantage of the Q filter appears to be the increased demand upon computation time. This can be severe, particularly for more complex models. For example, the two models treated here required nearly 50% more execution time than that without the sequential Q estimator. This aspect is substantial motivation for keeping the approximation structures relatively simple.

As a result of the application of the sequential Q estimator, we have shown that it is possible to improve the knowledge and confidence in the state estimates. It is most important, however, that the Q filter operation be initiated after the transient period of the state filter is passed.

5.7 Summary and Conclusions

In this chapter we have investigated the performance of model error and state noise covariance estimation algorithms as applied to the orbit determination problem of an SEP spacecraft subject to anomalies in the thrust program. We have established that some supplementary measurement type is useful in addition to the normal radar range rate observations. The use of the onboard star-vehicle-Earth angle, while providing such information, suffers two main disadvantages. First, high measurement accuracies are necessary to make the technique useful. Second, as the distance of the vehicle to the Earth increases, the measurement angle uncertainties translate into larger position uncertainties. This latter problem suggests some alternative approach, such as switching to the target body (in this case, the asteroid) as the approach phase is entered. This would provide the reverse effect with greater accuracy being obtained as the vehicle approaches the target. Another alternative is to employ a different measurement type. A potential candidate is quasi, very long baseline interferometry (QVLBI). This has been shown to offer greatly improved observation accuracies over that obtained with conventional range rate tracking (52), although additional development is required to make it practical.

The use of a state noise covariance matrix (Model 0) to maintain filter operation has been shown to be effective in preventing divergence of the estimates, and can yield accuracies of less than 500 km position error RSS. In general, the performance can be improved using any of the other

models investigated, although the second order models (2a and 2b) appear to offer greater accuracy and stability of the estimates. In particular, the superior performance (RSS ≈ 70 km) of Model 2b, being the closest in structure to the actual error, confirms the importance of accurate modeling to the maximum extent practical.

In all cases, a state noise covariance matrix is required in order to prevent filter saturation and subsequent divergence. However, with the estimation of the model errors, the state noise, and therefore the error covariances, are smaller than without model error estimation.

Application of the sequential Q estimator has been found to be a workable approach for estimating unknown state noise covariances. For the error approximation Model 1b, the Q filter greatly improved the RSS error, decreasing it from 450 km to about 120 km. The state RTC in both cases has been shown to be an adequate reflection of the confidence in the estimates, particularly in the case of Model 2b, where the RTC practically bounds the RSS. Of paramount importance in employing the state and Q filters is to insure that Q filter operation is not initiated until after the transient period of state filter operation. A potential problem is that of the choice of units for the dual state and Q filter computations. As reflected in Table 5.2, very small numerical values can be encountered for the Q error covariance matrix, S, and these must be compatible with the computational capabilities of the particular computer to be employed. Further, making the *a priori* S too large can result in estimates following the observation residuals, thus at least delaying accurate convergence of the estimates. Finally, the sequential Q estimator suffers the disadvantages of requiring up to 50% more computation time over that for just the state filter with *a priori* constant Q values.

In sum, we have shown that the thrust acceleration errors which, unchecked, produce extreme divergence of the state estimates, can be very adequately controlled. The techniques employed not only increase the navigation accuracy, but also knowledge of the actual dynamic model.

Chapter 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 Summary and Conclusions

This dissertation has been concerned with solving the fundamental problem of dynamic modeling errors in classical Kalman filtering. The very detrimental effects of such errors have been illustrated in both algebraic and numerical terms through the examples of the rendezvous and SEP vehicle estimation problems. In approaching the model error problem, various aspects and properties of the filtering elements have been presented, e.g., information, observability, etc., thus providing a fundamental base for the investigation. A number of non-adaptive algorithms have been presented which attempt to account for errors in the dynamic model. These range from the optimal approaches of employing a state noise covariance matrix and the more complex limited memory filter, to the suboptimal techniques of age-weighting, additive gain term, and gain scaling modifications. In all cases, these approaches compensate for the model error by simply keeping the gain at a high enough value to prevent saturation, and thus allow continued filter operation. There is, however, a certain minimum estimation accuracy obtainable, reflected in part by the steady state value of the error covariance, and determined by the corresponding filter parameters.

Except for the limited memory filter, adaptive forms of each of these error compensation techniques have been presented. Using the adaptive state noise covariance estimation algorithm due to Jazwinski, adaptive forms were derived for each of the other methods.

The powerful technique of adaptively estimating modeling errors has been presented, along with discussions of the assumed functional form or mathematical structure of these errors. A unifying derivation has been presented resulting in a relatively flexible structure suggested as a candidate function for use in model error compensation. Additionally, a new algorithm for sequentially estimating state noise covariances has been developed, based upon the ordinary Kalman filtering algorithm. In view of the non-negative definite property of the state noise covariance, and certain other assumptions, the algorithm produces constrained minimum variance estimates.

In applying the techniques to the rendezvous and SEP estimation problems, numerical results have offered some interesting conclusions. From the rendezvous problem, the performance of the ordinary gain scaling algorithm (b = constant) clearly indicates that this method is unacceptable to counter the effect of unbounded model errors. This is true for both the adaptive and non-adaptive forms. In their non-adaptive forms, the other algorithms have been shown to produce equivalent steady state performance when appropriate filter parameter values are employed. This is further substantiated by the adaptive formulations for which the performances are nearly the same. In fact, Jazwinski's state noise covariance estimation and the adaptive age-weighting algorithms produce identical performance. Further, the modified gain scaling and additive gain term methods in their adaptive forms also yield identical performance. Between the two sets, the former algorithms appear to offer slightly greater confidence in their estimates. The reason for this is not clear. However, one clue is that the former has filter parameters (Q and s) which can take on infinitely large values; the latter algorithms have parameters (α and β) ranging between zero and one.

In view of this, the conclusion suggested is that the differences may be due to numerical sensitivities, a condition brought on by finite computer word length. Thus, in the latter case, considerably more significant figures are required to achieve the same results as for Q and s.

Results obtained for the estimation of model errors, as applied to the rendezvous problem, clearly indicate the superior performance of this method. This is to be expected, since it is the only method which actually attempts to improve knowledge of the dynamic model. Coupled with an appropriately determined state noise covariance matrix, this technique is thus the best approach to the model error problem. For this reason, this approach was taken for investigating the solar electric propulsion mission. The accuracy of this approach in the SEP orbit determination study is demonstrated by the performance for the various assumed model structures. In general, higher order models are to be preferred since their structure tends to make them more adaptable to complex, as well as simple, error forms. This is substantiated by the more accurate estimation performance of the second order models. In view of the structure of the SEP acceleration error, and the rather remarkable performance of Model 2b, the importance of accurate but parexic modeling cannot be under-emphasized.

The sequential estimator for the state noise covariance matrix can be effective in maintaining operation of the Kalman filter, both with and without the formulation for estimating model errors, although it is rather demanding of computation time. For the rendezvous problem, the estimates of the state noise covariance were found to be somewhat erratic, indicating a rather acute sensitivity to the residuals. In the SEP mission simulation, the strategy of avoiding Q filter operation during the state filter transient periods proved to be the key to success in this application. Thus, performance

in the rendezvous problem could probably be improved using this philosophy. Further, in the rendezvous application, the T estimation utilized all *a posteriori* $\bar{\Delta}$ residuals, thus including the effects of transient state filter operation. For the SEP mission, the modification of the T estimator limiting it to a batch of 20 residuals appeared to be a more viable approach. In this configuration the Q filter operating in the state filter, post-transient period was found to be quite effective in improving state noise covariance values for Models 1b and 2b. The success obtained with this strategy suggests that the dual state-Q filter algorithm can be a useful estimation tool.

6.2 Recommendations for Further Investigation

One of the most apparent questions which remains unanswered is that of performance of the suboptimal adaptive and non-adaptive algorithms for the case of vector state variables. Some reflection will reveal that the algorithms could not be expected to be equivalenced in the same sense as for the scalar case, *viz.*, the rendezvous problem. For example, the age-weighting algorithm simply multiplies the error covariance by a scalar, thus scaling each element by the same amount. The additive gain term, however, adds different terms of varying value to the error covariance.

The additive gain term algorithm suffers a disadvantage in that the added term,

$$\beta R H^T / H H^T$$

affects the gain and error covariance only for those state elements which appear explicitly in the observation state relationship. As indicated in section 2.2, the gain scaling modification using

$$b = 1 + \frac{\beta R}{H \bar{P} H^T}$$

produces the gain expression of *(2.2.15)* which differs from that for the additive gain *(2.3.2)* by the presence of the error covariance, $\overline{P}$. As long as $\overline{P}$ is of full rank, the former gain,

$$K = (\overline{P}\,H^T + \beta\, R\,\overline{P}\,H^T/H\,\overline{P}\,H^T)(H\,\overline{P}\,H^T + R)^{-1} \qquad (2.3.15)$$

is altered for *all* terms in the state vector, regardless of their explicit appearance or absence in the measurement equation. Thus the original additive gain term cannot be successfully applied to, say, the model error estimation algorithm; the modified gain *(2.2.15)* can. In light of the fact that for the scalar problem the $\overline{P}$ divides out, this aspect was not investigated in the rendezvous problem. However, any vector problem (say, the rendezvous problem where range, ρ, is included with range rate, $\dot{\rho}$, to form a 2 vector) could be used to investigate this aspect.

Similar conditions occur for Jazwinski's Q estimation algorithm. Here the pseudo inverse,

$$\Lambda^{\#} = \frac{\Lambda^T}{\Lambda\,\Lambda^T} = \frac{\Gamma^T H^T}{H\,\Gamma\,\Gamma^T H^T}$$

usually suffers the same malady as the additive gain term. By employing a different pseudo inverse, say,

$$\Theta^{\#} = \frac{\overline{P}\,\Lambda^T}{\Lambda\,\overline{P}\,\Lambda^T}$$

state elements not in the measurement equation will still have corresponding non-zero state noise covariance elements.

In this study, the common structure given by the function of equation *(3.4.6)* has only been suggested for use in estimating model errors. It should be particularly interesting to investigate the ability of this common structure to adapt to different model error functions, both bounded and unbounded.

Considerable study of the sequential state noise estimator remains to be done, particularly with regard to sensitivity analyses. Some sensitivities have been examined with regard to various *a priori* values of S and T. However, additional study is warranted.

Finally, we reiterate that some of the algorithms may be combined to obtain hybrid techniques. The approaches here are many. One possibility is to apply the suggested modified gain scaling technique (section 2.2, with the gain given by *(2.2.15)*) to the model error estimation technique. This approach, while suboptimal, would require only the selection of the filter parameter, β, rather than three or more state noise covariance elements. Further, the adaptive form for estimating β could be employed, thereby making the algorithm completely adaptive.

Our concern in this study has been to investigate concepts and techniques for alleviating the dynamic model error problem in linear filtering. Many approaches have been discussed and analyzed, and advantages and disadvantages presented. The filtering and estimation process -- even in its most sophisticated forms -- can never be any better than the mathematical model representing the dynamic process. In the final analysis, then, there can be no substitute for accurate modeling, whether analytically determined or adaptively estimated.

Appendix A

MATRIX INVERSION LEMMA (SCHUR IDENTITY)

Define the positive definite matrix, X, as

$$X \equiv (A^T B A + C)^{-1} \tag{A.1}$$

Taken the inverse and pre-multiply by X to get

$$I = X A^T B A + X C \tag{A.2}$$

Post-multiply by C^{-1}:

$$C^{-1} = X A^T B A C^{-1} + X \tag{A.3}$$

or

$$X = C^{-1} - X A^T B A C^{-1} \tag{A.4}$$

Post-multiply *(A.3)* by $A^T B$ to obtain the following sequence:

$$\begin{aligned} C^{-1} A^T B &= X A^T B A C^{-1} A^T B + X A^T B \\ &= X A^T B(A C^{-1} A^T B + I) \\ &= X A^T B(A C^{-1} A^T + B^{-1}) B \end{aligned}$$

Post-multiply by B^{-1}:

$$C^{-1} A^T = X A^T B(A C^{-1} A^T + B^{-1})$$

or

$$X A^T B = C^{-1} A^T (A C^{-1} A^T + B^{-1})^{-1} \tag{A.5}$$

Substitute *(A.4)* into *(A.5)* and solve for X, thus obtaining an expression equivalent to *(A.1)*

$$X = C^{-1} - C^{-1} A^T (A C^{-1} A^T + B^{-1}) A C^{-1} \tag{A.6}$$

Appendix B

A AND B MATRIX ELEMENTS

A Matrix: Partition the A matrix into sub-matrices as follows:

$$A = \begin{bmatrix} 0 & I & 0 \\ A_{21} & 0 & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix}$$

These elements are given as

$I = 3 \times 3$ *identity matrix*

$A_{21} = [a_{ij}]$; $i = 4, 5, 6$; $j = 1, 2, 3$

$A_{23} = [a_{ij}]$; $i = 4, 5, 6$; $j = 7, 8, 9$

$A_{33} = [a_{ij}]$; $i, j = 7, 8, 9$

where

$$a_{41} = \frac{\mu}{|r|^3}\left[\frac{3X^2}{|r|^2} - 1\right] + \frac{XY}{|p|^3}\, e$$

$$a_{42} = \frac{\mu}{|r|^3}\left[\frac{3XY}{|r|^2}\right] - \frac{1}{|p|}\left[1 - \frac{Y^2}{|p|^2}\, e\right]$$

$$a_{43} = \frac{\mu}{|r|^3}\left[\frac{3XZ}{|r|^2}\right]$$

$$a_{51} = \frac{\mu}{|r|^3}\left[\frac{3YX}{|r|^2}\right] + \frac{1}{|p|}\left[1 - \frac{Y^2}{|p|^2}\, e\right]$$

$$a_{52} = \frac{\mu}{|r|^3}\left[\frac{3Y^2}{|r|^2} - 1\right] - \frac{XY}{|p|^3}\, e$$

$$a_{53} = \frac{\mu}{|r|^3}\left[\frac{3YZ}{|r|^2}\right]$$

$$a_{61} = \frac{\mu}{|r|^3}\left[\frac{3YZ}{|r|^2}\right]$$

$$a_{62} = \frac{\mu}{|r|^3}\left[\frac{3ZY}{|r|^2}\right]$$

$$a_{63} = \frac{\mu}{|r|^3}\left[\frac{3Z^2}{|r|^2} - 1\right]$$

$a_{47} = -Y/|p|$, $a_{57} = X/|p|$, $a_{ij} = 0$ *for other* i, j

e is the estimated thrust acceleration error, and $|p| = (X^2 + Y^2)^{\frac{1}{2}}$

Model 1a: $a_{ij} = 0$, $i,j = 7, 8, 9$

Model 1b: $a_{77} = -\alpha$, $a_{78} = -e$, $a_{ij} = 0$ *for other* i,j

Model 2a: $a_{78} = 1$, $a_{ij} = 0$ *for other* i,j

Model 2b: $a_{78} = 1$, $a_{87} = -\beta$, $a_{89} = -e$, $a_{ij} = 0$ *for other* i,j

For Model 0, $e = 0$ in A_{21}, and $A_{23} = A_{33} = 0$.

<u>B Matrix</u>: The B matrix is defined in partitioned form as

$$B = \begin{bmatrix} I & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & I \end{bmatrix}$$

where

$$M = \begin{bmatrix} X/|p| & -Y/|p| & 0 \\ Y/|p| & X/|p| & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which is the transformation matrix from the orbital frame to the heliocentric frame.

Appendix C

OBSERVATION - STATE RELATION PARTIAL DERIVATIVES (H MATRIX)

Range Rate Measurements, $\dot{\rho}$:

$$\frac{\partial \dot{\rho}}{\partial X_i} = \left[(\dot{X}_i - \dot{X}_{si}) - (X_i - X_{si}) \, (\dot{\rho}/\rho) \right] / \rho \ ,$$

$$\frac{\partial \dot{\rho}}{\partial \dot{X}_i} = (X_i - X_{si}) / \rho \quad , \quad i = 1, 2, 3$$

where X_i and $\dot{X}_i$ are the heliocentric position and velocity components of the vehicle; and X_{si} and $\dot{X}_{si}$ are the heliocentric position and velocity components of the tracking station. The derivatives with respect to the other state elements (depending upon the approximating model) are zero.

Star-Vehicle-Earth Angle, ξ:

$$\frac{\partial \xi}{\partial \dot{X}_i} = \left[s_{Xi} - \cos \xi \, \frac{(X_{ei} - X_i)}{|R_e - r|} \right] / (|R_e - r|^2 \sin \xi), \quad i = 1, 2, 3$$

where the X_{ei} are the heliocentric position components of the Earth; the s_{Xi} are the navigation star unit vector heliocentric components; R_e and r are the Earth and vehicle heliocentric position vectors. The rest of the partial derivatives are zero.

Appendix D

THE ERGODIC PROCESS

In a process which is ergodic in the mean, theory predicts 68.27% for normally distributed errors (46, 98). If the process is ergodic in the mean, the time sample averages are representative of the ensemble sample averages, that is averaging more and more values along the time axis results in convergence to the ensemble mean. A necessary and sufficient condition that a process, $x(t)$, $0 \leq t < \infty$, be ergodic in the mean is that it be wide-sense or weakly stationary, and that

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T E\{[x(t) - \mu]^2\} \, d\, t = 0$$

or in the discrete case,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N-1} E\{[x - \mu]^2\} = 0$$

See, for example, (46, 166-167). The assumption of ergodicity in the mean for the range rate error process, $\tilde{\dot{\rho}}(t)$, is reasonable as long as the process is stationary and the time average converges to $E\{\tilde{\dot{\rho}}\} = 0$, a condition guaranteed if the limit equations above are fulfilled. In our investigation, this is not strictly the case; however, for short time periods it is approximately true. Further, for a given mean and covariance, the Gaussian distribution represents the maximum uncertainty, entropy, (47, 613). Thus the comparison of the time sample with theoretical normal (Gaussian) distribution limits is reasonable as long as the ergodicity assumption is justifiable.

BIBLIOGRAPHY

LITERATURE CITED

1. Gauss, Karl F., *Theory of the Motion of the Heavenly Bodies Moving About the Sun in Conic Sections*. New York: Dover Publications, Inc., 1963 (reprint).

2. Legendre, A. M., *Nouvelles méthodes pour la détermination des orbites des cométes*. Paris: 1806.

3. Fisher, R. A., "Theory of Statistical Estimation," Proc. Cambridge Philos. Soc., vol. 22, 1925, 700.

4. Kalman, R. E., and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," J. Basic Eng., Trans. ASME, Vol. 83D, pp. 95-108, 1961.

5. Jazwinski, A. H., *Stochastic Processes and Control Theory*, New York: Academic Press, 1970.

6. Kalman, R. E., "On the General Theory of Control Systems," Proc. 1st Intern. Congr. Autom. Control, London, vol. 1, p. 481, 1961.

7. Sorenson, H. W., "On the Error Behavior in Linear Minimum Variance Estimation Problems," IEEE Trans. Autom. Control, vol. AC-12, pp. 557-562, 1967.

8. Leondes, C. T., (ed.), *Theory and Applications of Kalman Filtering (AGARDograph 139)*. London: Technical Editing and Reproduction Ltd., 1970. (Reproduced by National Technical Information Service, Springfield, Va.)

9. The Martin Company, *Orbital Flight Handbook, Volume I, Part 2 - Mission Sequencing Problems*. pp. VII-13 - VII-54.

10. Heffes, H., "The Effect of Erroneous Models on the Kalman Filter Responses," IEEE Trans. Autom. Control, July, pp. 541-545, 1966.

11. Schlee, F. H., C. J. Standish, and N. F. Toda, "Divergence in the Kalman Filter," AIAA J., Vol. 5, pp. 1114-1120, 1967.

12. Price, C. F., "An Analysis of the Divergence Problem in the Kalman Filter," IEEE Trans. Autom. Control, December, pp. 699-702, 1968.

13. Huddle, J. R., and D. A. Wismer, "Degradation of Linear Filter Performance Due to Modeling Error," IEEE Trans. Autom. Control, August, pp. 421-423.

14. Neal, S. R., "Linear Estimation in the Presence of Errors in the Assumed Plant Dynamics," IEEE Trans. Autom. Control, October, pp. 592-594, 1967.

15. Meditch, J. S., *Stochastic Optimal Linear Estimation and Control.* New York: McGraw-Hill Book Company, 1969.

16. Tarn, T. J., and J. Zaborszky, "A Practical, Nondiverging Filter," AIAA J., vol. 8, pp. 1127-1133, 1970.

17. Fagin, S. L., "Recursive Linear Regression Theory, Optimal Filter Theory, and Error Analysis of Optimal Systems," IEEE Intern. Conv. Record, vol. 12, pp. 216-240, 1964.

18. Miller, R. W., "Asymptotic Behavior of the Kalman Filter with Exponential Aging," AIAA J., vol. 9, pp. 537-539, 1971.

19. Schmidt, S. F., "Estimation of State with Acceptable Accuracy Constraints," Analytical Mechanics Associates, Inc., Interim Report No. 67-4, January, 1967.

20. _____, "Compensation for Modeling Errors in Orbit Determination Problems," Analytical Mechanics Associates, Inc., Interim Report No. 67-16, November, 1967.

21. Jazwinski, A. H., "Limited Memory Optimal Filtering," Analytical Mechanics Associates, Inc., Interim Report No. 67-11, August, 1967.

22. Mehra, R. K., "Approaches to Adaptive Filtering," Proc. 9th IEEE Symposium on Adaptive Processes: Decision and Control, The University of Texas at Austin, December, 1970.

23. Kailath, T., "An Innovations Approach to Least-Squares Estimation, Part I: Linear Filtering in Additive White Noise," IEEE Trans. Autom. Control, Vol. AC-13, pp. 645-655, December, 1968.

24. Jazwinski, A. H., "Adaptive Filtering," Proc. IFAC Symp. Multivariable Control Systems, Dusseldorf, Germany, Vol. 2, pp. 1-15.

25. Magill, D. T., "Optimal Adaptive Estimation of Samples Stochastic Processes," IEEE Trans. Autom. Control, Vol. AC-10, pp. 434-439, 1965.

26. Abramson, Jr., P. D., "Simultaneous Estimation of the State and Noise Statistics in Linear Dynamical Systems," Experimental Astronomy Laboratory, Massachusetts Institute of Technology, Report No. TE-25, 1968.

27. Sage, A. P., and G. W. Husa, "Adaptive Filtering with Unknown Prior Statistics," Proc. Joint Autom. Control Conf., 1969.

28. Levy, L. J., "Adaptive Estimation Algorithms," Engineering Research Institute, Iowa State University, Tech. Report ERI-93400, December, 1970.

29. Mehra, R. K., "On the Identification of Variances and Adaptive Kalman Filtering," IEEE Trans. Autom. Control, vol. AC-15, pp. 175-184, 1970.

30. _____, "On-Line Identification of Linear Dynamic Systems with Applications to Kalman Filtering," IEEE Trans. Autom. Control, vol. AC-16, pp. 12-21, 1971.

31. Jazwinski, A. H., and A. E. Bailic, "Adaptive Filtering," Analytical Mechanics Associates, Inc., Interim Report No. 67-6, March, 1967.

32. Ingram, D. S., "Orbit Determination in the Presence of Unmodeled Accelerations," Applied Mechanics Research Laboratory, The University of Texas at Austin, Report No. AMRL-1022, January, 1971.

33. Tapley, B. D., and B. E. Schutz, "Estimation of Unmodeled Forces on a Lunar Satellite," 23rd Cong. of the Int. Astronautical Fed., Vienna, Austria, Oct. 8-15, 1972.

34. Tapley, B. D., and H. Hagar, Jr., "Navigation Strategy and Filter Design for Solar Electric Missions," Applied Mechanics Research Laboratory, The University of Texas at Austin, Report No. AMRL-1040, May, 1972.

35. _____, "Estimation of Unmodeled Thrusting Forces on a Low-Thrust Space Vehicle," Proc. Third Symp. on Nonlinear Estimation Theory, San Diego, California, September, 1972.

36. Russell, R. K., and D. W. Curkendall, "On Modeling Continuous Accelerations as Piecewise Constant Functions," Jet Propulsion Laboratories, JPL Technical Report 32-1526, vol. VIII, pp. 45-52.

37. Carpenter, G. C., and E. T. Pitkin, "Orbit Determination for a Thrusted Space Vehicle," AIAA/AAS Astrodynamics Conference, Princeton, New Jersey, AIAA Paper No. 69-901, August, 1969.

38. Lainiotis, D. G., "Optimal Adaptive Estimation: Structure and Parameter Adaptation," IEEE Trans. Autom. Control, vol. AC-16, pp. 160-170, 1971.

39. Liebelt, P. B., *An Introduction to Optimal Estimation*. Reading: Addison-Wesley Publishing Co., 1967.

40. Papoulis, A., *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill Book Co., 1965.

41. Tapley, B. D., and D. S. Ingram, "Orbit Determination in the Presence of Unmodeled Accelerations," Proc. Second Symp. Nonlinear Estimation Theory, San Diego, California, September, 1971.

42. Wiley, C. R., Jr., *Advanced Engineering Mathematics*. New York: McGraw-Hill Book Co., 1966.

43. Tapley, B. D., and G. H. Born, "Sequential Estimation of the State and the Observation-Error Covariance Matrix," AIAA J., vol. 9, no. 2, pp. 212-217, 1971.

44. Bryson, A. E., and Y. C. Ho, *Applied Optimal Control*. Waltham: Blaisdell Publishing Co., 1969.

45. Deutsch, Ralph, *Estimation Theory*. Englewood Cliffs: Prentice-Hall, Inc., 1965.

46. Gray, H. L., and P. L. Odell, *Probability for Practicing Engineers*. New York: Barnes & Noble, Inc. 1970.

47. Korn, G. A., and T. M. Korn. *Mathematical Handbook for Scientists and Engineers*. New York: McGraw-Hill Book Co., 1968.

48. Cheney, E. W., *Introduction to Approximation Theory*. New York: McGraw-Hill Book Company, 1966.

49. Jordan, J. F., G. A. Madrid, and G. E. Pease, "The Effects of Major Error Sources in Planetary Spacecraft Navigation Accuracies," *Journal of Spacecraft and Rockets*, vol. 8, no. 9, pp. 920-926, September, 1971.

50. Rourke, K. H., and J. F. Jordan, "Guidance and Navigation for Solar Electric Interplanetary Missions," AIAA 8th Electric Propulsion Conference, Stanford, California, AIAA Paper No. 70-1152, August, 1970.

51. Curkendall, D. W., and R. R. Stephenson, "Earthbased Tracking and Orbit Determination -- Backbone of the Planetary Navigation System," *Astronautics and Aeronautics*, pp. 30-36, May 1970.

52. Ondrasik, V. J., and K. H. Rourke, "Applications of Quasi-VLBI Tracking Data Types to the Zero Declination and Process Noise Problems," AAS/AIAA Astrodynamics Specialists Conference, Ft. Lauderdale, Florida, August 1971.

ADDITIONAL REFERENCES

1. Andrews, A., "A Square Root Formulation of the Kalman Covariance Equations," *AIAA Journal*, vol. 6, pp. 1165-1166, June, 1968.

2. Deyst, John J., Jr., and Charles F. Price, "Conditions for Asymptotic Stability of the Discrete Minimum Variance Linear Estimator," *IEEE Transactions on Automatic Control*, pp. 702-705, December, 1968.

3. Duren, D. J., "Convergence of the Difference Equation for the Error Covariance Matrix Arising in Kalman Filter Theory," Tech Report ERI-94900, Engineering Research Institute, Iowa State University, 1971.

4. Griffin, Robert E., and Andrew P. Sage, "Large and Small Scale Sensitivity Analysis of Optimum Estimation Algorithms," *IEEE Transactions on Automatic Control*, pp. 320-328, August, 1968.

5. Jazwinski, Andrew H., "Adaptive Filtering in Satellite Orbit Estimation," presented at the Second Asilomar Conference on Circuits and Systems, Pacific Grove, California, October 30, 1968.

6. LeMay, Joseph L., and Allen P. Stubberud, "The Time Correlation of Estimation Errors with Aerospace Applications," presented at the AIAA Guidance, Control, and Flight Mechanics Conference, University of California, Santa Barbara, August 17-19, 1970.

7. Mann, Lawrence, Jr., *Applied Engineering Statistics for Practicing Engineers*, Barnes and Noble, Inc., New York, 1970.

8. The Martin Company, Space Systems Division, *Orbital Flight Handbook, Vol. I, Part 1 - Basic Techniques and Data*, National Aeronautics and Space Administration, Washington, D.C., 1963.

9. McRary, John W., "Detection of Unmodeled Errors in Recursive Trajectory Estimation," *AIAA Journal*, vol. 7, No. 9, pp. 1790-1792, September, 1969.

10. Nahi, Nasser E., and Brian M. Schaffer, "Decision-Directed Adaptive Recursive Estimators: Divergence Prevention," *IEEE Transactions on Automatic Control*, vol. AC-17, No. 1, February, 1972.

11. Potter, James E., and Robert G. Stern, "Statistical Filtering of Space Navigation Measurements," Experimental Astronomy Laboratory, Massachusetts Institute of Technology, report # RE-3, August, 1963.

12. Sage, Andrew P., *Optimum Systems Control*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1968.

13. Schmidt, Stanley F., "Compensation for Modeling Errors in Orbit Determination Problems," Report No. 67-16, Analytical Mechanics Associates, November, 1967.

14. Weiss, I. M., "A Survey of Discrete Kalman Bucy Filtering with Unknown Noise Covariances," AIAA paper No. 70-955, presented at the AIAA Guidance, Control, and Flight Mechanics Conference, Santa Barbara, California, August 17-19, 1970.

15. Wiener, Norbert, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, The M.I.T. Press, Massachusetts Institute of Technology, Cambridge, Massachusetts, August, 1949.